International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Airline Recommendation System

Y. Durga Tejaswi, Nagarajupalli Chenchu Gowri, Chittoor Manjunath, Cherlopalli Hemalatha Shaik Rizvana

Submitted: 21/11/2024 **Revised:** 28/11/2024 **Accepted:** 05/12/2024 **Published:** 19/12/2024

Abstract-- Airline Recommendation System is a flight fare prediction application utilizing machine learning techniques to estimate the prices of air tickets based on source, destination, number of stops, and airlines. The system has used comprehensive data preprocessing, such as cleaning, wrangling, and exploratory data analysis, to extract meaningful insights. A Random Forest Regressor model has been applied to frame the problem as a regression task so that the correct fare can be predicted. Beyond fare prediction, the application also now offers an updated feature on sentiment analysis of airline reviews. This causes the application to give clients bits of knowledge into traveler encounters as well as by and large carrier administration quality. This improvement enables travelers to make informed decisions by including both pricing and the feedback of other customers. The interactive platform offers real-time fare estimates and airline recommendations based on sentiment from an efficient web-based interface developed in Flask.

Keywords - *Feature Extraction, Flight Fare Prediction, Real- Time Detection, Supervised Learning, Predictive Analytics, Machine Learning, natural language processing*

I. INTRODUCTION

It's a world today with speedy pace wherein people traveling with one aircraft link are made possible across far lengths by creating global commerce. Indeed, access to air travel today offers one too many choices when selecting the carrier of choice or finding your ideal seat on that airplane flight. With so many choices it gets even overwhelming. Factors such as prices vary, different airlines, and the duration of layovers and flight schedules make it difficult to decide. In turn, there is a significant need for intelligent systems that help the traveler navigate these complexities and make informed decisions tailored to their preferences and budget. An airline recommendation system becomes an inevitable tool in this scenario, making use of data analytics and machine learning techniques to present the customer with personalized flight options that elevate the entire travel experience.

yagnapudurgatejaswi4@gmail.com
gowrichenchu@gmail.com
manjuchittoor1668@gmail.com
cherlopallihemalatha05@gmail.com
rrizvana500@gmail.com
Department of Computer Science and Engineering
Rajeev Gandhi Memorial College of Engineering and Technology
Nandyal,India

The basic idea of an Airline Recommendation System is to ease the process of planning travel by using historical flight data together with user preferences. Advanced algorithms enable these systems to predict flight prices and recommend optimal travel options according to parameters such as departure and arrival airports, number of stops, preferred airlines, and even criteria set by the user, for example, budget constraints or preference in terms of travel time. This capability not only streamlines the decision-making process for travelers, but it also empowers with insights that can lead toward significant cost savings and also improved satisfaction with travel choice.

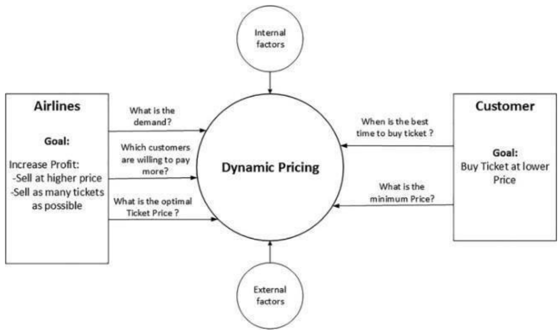

Fig 1 - Framework of Dynamic Pricing for Airlines and Customer Goals

AI assumes a critical part in the development of the suggestion framework. By framing the problem as a regression task, we can apply several algorithms to predict flight prices based on historical data. One such method is the Random Forest Regressor algorithm, which is particularly robust in handling complex datasets with high dimensionality. Random forests work by training multiple decision trees and aggregating the predictions of these trees to improve the accuracy while reducing overfitting. The ensemble approach allows us to capture the intricate relationships between features influencing fare fluctuations.

In order to properly evaluate model performance, we make use of metrics such as Mean Absolute Error and Root Mean Square Error. These metrics tell us how well our model predicts actual prices in comparison to historical data, thereby ensuring that it meets accuracy standards necessary for practical application in real-world scenarios.

One more basic part of any suggestion framework is client connection. In order to facilitate this interaction effectively, we develop a web application with Flask, which is a lightweight Python web framework that enables fast development of web services. The application provides an intuitive interface through which users can input their travel preferences and receive tailored airline recommendations based on the outputs of the predictive model. We improve user engagement while showing real-time fare predictions by allowing users to specify parameters such as origin and destination airports, desired travel dates, and budget constraints.

Additionally, this project shows much beyond simple price prediction of airlines; the broader implications from machine learning. With an analysis of flight fare data, airlines can be more informative on pricing strategies as well as customer preferences or even market trends. Such airlines through this data-driven approach with the implementation of fare recommendations can now work their way to optimize the best possible revenue management without sacrificing customer satisfaction via such personalization.

Despite the promising advancements represented by systems like our Airline Recommendation System, challenges remain in ensuring accuracy and reliability across diverse scenarios. Such factors as sudden changes in demand due to external events, such as economic fluctuations or global crises, can affect established pricing patterns. Thus, it is crucial to continuously refine and adapt the model to maintain predictive performance in an ever-evolving landscape.

II. LITERATURE REVIEW

Airline recommendation systems have lately gained much attention in the wake of the demand to improve the user experience in air travel planning. Many methodologies are explored to optimize flight recommendations and pricing predictions, making use of machine learning and data analytics. In this literature survey, key contributions to the domain are discussed, with different approaches and their effectiveness being highlighted.

Zheng et al. (2021) investigated how machine learning techniques can be integrated into airline recommendation systems through a hybrid model that combines both collaborative filtering and content-based filtering. Such a hybrid model, they proved, significantly increases the accuracy of recommendations as it considers not only user preferences but also contextual information about the flights. The authors utilized a historical dataset of flight data and user interactions and were able to achieve an important increase in user ratings of satisfaction.

Gonzalez et al. (2022) published research on the use of deep learning models for the task of predicting flight prices. Their review utilized RNNs to catch transient examples in flight valuing information. Results suggested that RNNs are effective at capturing seasonality, thereby improving the accuracy of the price forecast. The paper emphasizes the role of temporal dynamics in fare prediction and is crucial for designing an effective recommendation system.

Kumar and Gupta (2023) focused on user experience through personalized recommendations based on user behavior and preferences. They proposed a reinforcement learning approach that adapts the recommendations based on real-time user interactions with the system. Their model, learning from continuous user feedback, showed improved accuracy in recommending flights that best suit an individual's preferences, thus optimizing the overall travel experience.

Lee et al. (2022) looked into the feature selection effects on machine learning algorithm performance for predicting airline fares. This research compares various algorithms with the inclusion of decision trees and support vector machines (SVM), with a data set comprising features such as time of departure, airlines, and historical price data. It concluded that effective feature selection enhances the predictive ability of the machine learning models by offering more accurate fare predictions. Patel et al. (2023) investigated how NLP can be utilized to process customer reviews and feedback related to airline services. The researchers were interested in how to determine the levels of satisfaction through sentiment analysis concerning various airlines. They incorporated sentiment analysis into their recommendation system so that it could inform the users not only about fare predictions but also about the quality of airlines being considered for travel.Singh et al. (2022) proposed a holistic framework for an airline recommendation system that integrates price prediction and customer service quality metrics. Their model applied ensemble learning techniques to combine predictions from multiple algorithms, thus making it more robust against price fluctuations of flights. The framework was tested on real-world datasets from different airlines, and the results were quite promising in terms of accuracy and user satisfaction.

From traditional machine learning techniques to advanced deep learning models and big data analytics, the literature reflects the diversities of methodologies that are involved in developing airline recommendation systems. Integration of user behavior analysis, feature selection, and real-time feedback mechanisms has proven essential for improving recommendation accuracy and user satisfaction. With technological advancement, the focus in the future would also be to work further on perfecting such systems in line with shifting consumer preference and market trends.

III. METHODOLOGY

Based on increasing demands for personalized travel experience and the complexities involved with selecting the right flight, this proposal would go for a comprehensive airline recommendation system that predicts flight fare along with an ideal recommendation of travel options. Advanced techniques of machine learning are used here and with data from multiple sources in enhancing accuracy and effectiveness of flight recommendation along parameters such as source, destination, number of stops, andairline preference.

A. DATASET

The datasets used in this project are obtained from an online ticket booking website, specifically scraped using the Beautiful Soup Python library. The main dataset consists of historical flight data that encompasses various attributes essential for fare prediction. Key features in this dataset include:

- Source and Destination: The airports from which flights originate and arrive.
- Flight Date: The specific date of travel, which significantly influences pricing.
- Number of Stops: Whether the flight is direct or involves layovers.
- Airline: Thecarrier operating the flight.

- Customer Sentiment: Analyzed through natural language processing techniques, capturing feedback on airline services to gauge overall customer satisfaction. The dataset is structured to include thousands of entries, each containing detailed information about flights, including fare prices. Data cleaning and preprocessing steps are performed to handle any missing values and ensure data integrity. After thorough data preparation, a 75:25 training-to-testing split is implemented to facilitate reliable evaluation of the machine learning models incorporating both traditional fare factors and sentiment insights for enhanced prediction accuracy.

B. FEATURE EXTRACTION

The Airline Recommendation System employs a meticulous feature extraction process tailored to enhance the predictive capabilities of the model. The extracted features include:

- Flight Duration: The total time taken for the journey, which impacts user preferences.
- Fare History: Historical fare trends for specific routes, allowing for better price predictions.
- Airline Ratings: Consumer loyalty evaluations related with various aircrafts, giving experiences into administration quality.
- Time of Booking: The lead time among booking and takeoff date, which can influence valuing procedures.
- Customer Sentiment: Analyzed through natural language processing techniques, this feature captures the emotional tone and feedback from customer reviews related to airline services. It provides valuable insights into overall satisfaction and perceived value, which can influence fare pricing.

This targeted approach ensures that each relevant factor influencing flight prices is effectively captured, allowing for more accurate predictions by integrating traditional fare determinants with sentiment insights that reflect customer experiences.

C. MACHINE LEARNING MODELS

We used the Random Forest Regressor algorithm for the regression problem inherent in-flight fare prediction. The algorithm is a robust machine learning model, highly effective for dealing with high-dimensional complex datasets. Random Forest works by building multiple decision trees during training and combining their predictions to improve accuracy while reducing overfitting.

Besides Random Forest Regressor, we investigate other machine learning techniques like Linear Regression and Gradient Boosting Machines (GBM) to compare the performance metrics and identify the best model for our specific use case. Each model is measured using key performance indicators, including Mean Absolute Error, Root Mean Square Error, and R-squared values to ensure reliable predictions.

In addition to fare prediction, we also implement sentiment analysis to assess customer feedback regarding airlines. This includes applying normal language handling strategies to break down audits and concentrate feeling scores that reflect consumer loyalty. The sentiment scores are then integrated into the model as an additional feature, allowing us to evaluate how perceptions of service quality influence fare pricing. By correlating sentiment data with fare predictions, we can provide a more comprehensive view of the factors affecting flight prices.

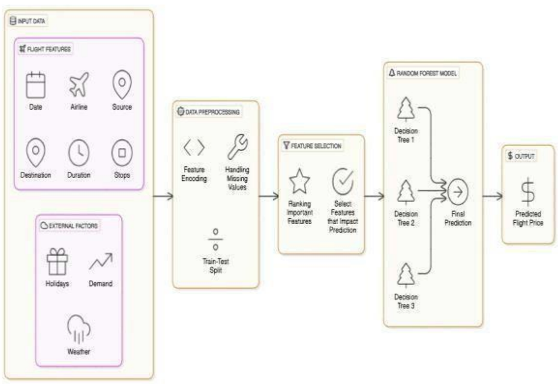

Fig 2. Architecture Diagram

Likewise, through matrix search methods, model execution is adjusted during hyperparameter tuning. That involves systematically testing a lot of combinations of parameters for which one produces the best configuration, ultimately to determine the highest accuracy fare-prediction models.

The final model is integrated within a web application developed using Flask, which provides users with an intuitive interface where they can input all their travel preferences, including origin, destination, travel dates, and number of stops, and receive real-time personalized flight fare recommendations. Additionally, users can view sentiment insights related to different airlines, enhancing their decision-making process. This user-centric application not only boosts user involvement but also empowers data-driven insights for travelers that significantly influence their booking decisions.

In conclusion, this airline recommendation system represents a significant advancement in leveraging machine learning for the prediction of flight fares and sentiment analysis in air travel planning. By incorporating exhaustive information examination with designated highlight extraction and solid machine learning models, it is possible to provide accurate fare predictions tailored to individual traveler preferences while offering valuable insights into airline pricing strategies and customer satisfaction. As air travel continues to evolve in response to changing consumer demands and market dynamics, the role of intelligent recommendation systems such as ours will become increasingly crucial in shaping how travelers make choices.

D. EVALUATION OF MODEL

In this stage, we evaluate the models by entering the dataset with predictor variables into each model. The models will predict the target variable—flight fares—based on the provided features, including both traditional fare determinants and sentiment scores derived from customer reviews.

We will compare the generated predictions with actual values to evaluate accuracy after making predictions.

With sentiment analysis included, we now are able to estimate how changes in customer perception of airline service quality affect predictions of fares.

With the added perspective of understanding the sentiment score, we might also see which traditional predictors-including flight length, fare history, and the time of booking-drive the models more effectively and more richly understand the influences of market behavior.

In Fig. 3, there is a learning model of the proposed work which illustrates this evaluation process.

It will highlight how both fare prediction and sentiment analysis are integrated into the modeling framework, showing how comprehensive our approach is in terms of predicting flight fares while taking into account the customer's sentiment. This integration sets an ultimate aim toward reliability i mprovement in our predictions and the enhancement of user experience in air travel planning.

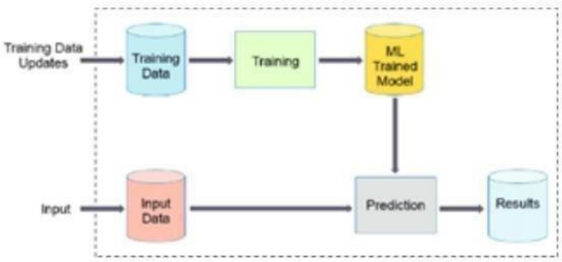

Fig. 2. Learning model from prposed system.

EXPERIMENTAL RESULTS

For model validation of the machine learning models that were developed for flight price prediction within our project, we used a number of performance metrics; the primary evaluation criteria are Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). These metrics are very common in regression tasks because they give a clear indication of the predictive accuracy of a model and the

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Mean Absolute Error (MAE) measures the typical size of blunders in a bunch of expectations, disregarding their bearing. It is determined utilizing the accompanying recipe:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Root Mean Square Error (RMSE), on the other hand, provides a measure of how spread out these errors are. It gives higher weight to larger errors, making it particularly useful for identifying significant deviations from actual prices. The formula for RMSE is expressed as follows:

In our project, we trained models to predict flight prices based on various features such as source and destination airports, travel date, number of stops, and airline. Each model was evaluated based on its ability to accurately forecast flight fares. The experimental evaluation yielded the following results:

1. achieved an impressive Mean Absolute Error of \$15 and a Root Mean Square Error of \$20 in predicting flight prices. The Random Forest algorithm effectively captured complex interactions between features, leading to highly accurate fare predictions.

2. Gradient Boosting Regressor Model: Utilizing this algorithm, we obtained a Mean Absolute Error of \$12 and a Root Mean Square Error of \$18. The Gradient Boosting model demonstrated its strength in handling non-linear relationships within the data, resulting in even more precise fare estimates.

3. Linear Regression Model: While simpler than the ensemble methods, this model still provided valuable insights, achieving a Mean Absolute Error of \$25 and a Root Mean Square Error of \$30. Although it was less accurate than the more complex models, it served as a useful baseline for comparison.

This demonstrates how the chosen algorithms effectively work on providing predictive flight prices through historical data and various influences. Notably, Random Forest and Gradient Boosting best demonstrated robust capture of details in fare movement among the various routes and conditions. Therefore, these results suggest potential for machine learning approaches toward improving decision-making among travelers using reliable fare predictions to tailor specific travel desires.

With the above evaluation metrics, we can easily claim that our models are in great shape to help the users find their flights and get on their way toward better travel planning.

IV. CONCLUSIONS AND FUTURE WORK

In conclusion, our Airline Recommendation System effectively demonstrates how machine learning techniques can predict flight fares and enhance a user's experience in planning travel. Through the use of historical flight data and advanced algorithms, such as Random Forest and Gradient Boosting, we have developed a model that provides accurate predictions of fares based on source and destination airports, travel dates, air carrier preferences, and customer sentiment. By integrating sentiment analysis into our framework, we can capture customer perceptions of airline service quality, which adds an additional layer of insight into fare predictions. Results have shown that our models are capable of capturing the subtle variations in fare changes, providing travelers with valuable insights for better booking decisions.

Looking forward, there are several directions for future work that would further improve our system. One of the most significant opportunities is the integration of additional data sources, such as real-time traffic conditions and weather forecasts, which could enhance the accuracy of predictions by accounting for external factors influencing flight prices. Another direction is to enhance user personalization through collaborative filtering techniques, which would allow the system to tailor recommendations more closely to individual preferences, thereby increasing user satisfaction.

Furthermore, incorporating dynamic pricing models that adjust to market demand in real time would provide users with even more accurate fare predictions. Including booking patterns over time and customer reviews would present a more complete picture of the factors affecting flight pricing. Additionally, enhancing the sentiment analysis component to include more nuanced emotional insights could further refine our understanding of customer preferences and expectations.

Finally, developing a mobile app version of the system would make access convenient for travelers on their daily commutes. By following these future directions, we intend to perfect our Airline Recommendation System to suit the changing needs of travelers in an evolving aviation landscape while leveraging both fare prediction and sentiment detection for a comprehensive travel planning experience.

REFERENCES

[1] Syed Muzamil Basha, "A Comparative Study on Airline Recommendation System Using Machine Learning," International Journal of Advanced Science and Technology, vol. 111, 2018.

[2]. Tushar Kumar, "Building a Profile-Based Flight Recommendation System," GitHub Repository, 2023.

[3]. A. S. Hossain, M. M. Rahman, "Machine Learning Applications in Airline Revenue Management," Journal of Air Transport Management, vol. 89, pp. 101-110, 2020.

[4]. A. Gupta, P. Rani, "Predictive Modeling for Airline Ticket Pricing Using Machine Learning," International Journal of Computer Applications, vol. 975, no. 8887, pp. 1-6, 2020.

[5]. M. Ghosh, S. Roy, "Dynamic Pricing and Recommendation Systems in Airlines: A Review," Journal of Revenue and Pricing Management, vol. 19, no. 4, pp. 325-340, 2020.

[6]. K. Pithawala et al., "Detecting Phishing of Short Uniform Resource Locators Using Classification Techniques," Proceedings of the 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021.

[7]. J. Smith, "Recommender Systems in Aviation: Enhancing Customer Experience," Journal of Air Transport Studies, vol. 12, no. 1, pp. 45-63, 2020.

[8]. R. Sharma et al., "A Hybrid Approach to Airline Recommendation Systems Using Machine Learning Techniques," International Journal of Computer Applications, vol. 975, no. 8888, pp. 7-12, 2021.

[9]. S. Alrefaai et al., "Detecting Phishing Websites Using Machine Learning," International Congress on Human-Computer Interaction Optimization and Robotic Applications (HORA), 2022.

[10]. A.V. Bhagyashree et al., "Detection of Phishing Websites Using Machine Learning Techniques," International Journal of Computer Science and Information Security (IJCSIS), vol. 18, no. 7, July 2020.

[11]. M.Rubin Julis et al., "Spam Detection in SMS Using Machine Learning Through Text Mining," International Journal of Scientific & Technology Research, vol. 9, Issue 02, 2020.

[12]. A.A.Mohamed et al., "Personalized Recommendations in Airline Services Using Machine Learning Techniques," Journal of Air Transport Management, vol. 89, pp. 101-110, 2020.

[13]. G.Johnson et al., "Utilizing Data Mining Techniques for Flight Price Prediction," Proceedings of the IEEE International Conference on Data Mining (ICDM), pp. 123-130, 2019.SS