International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Intelligent Fare Dynamics Modeling Using Machine Learning for Adaptive Airline Pricing Forecasts

SK Fhysuddin[1*], Pallam Tejasri[1], Avula Swapna Sri[1], Thota Sai Spurthi[1], Dasari Bharath Kishore[1], Bommanaboina Purnima[1]
[1]Department of Electronics and Communication Engineering, Mother Teresa Institute of Science and Technology, Sanketika Nagar, Kothuru, Sathupally, Khammam, 507303, Telangana, India
*Correspondence: SK Fhysuddin

ABSTRACT
This research addressed the problem of predicting flight ticket prices by analyzing various factors that influence fare variations. Key attributes such as departure time, destination, flight duration, and seasonal demand were considered, as these parameters significantly affect pricing. Flight fares are highly dynamic and tend to fluctuate based on travel schedules, route characteristics, and peak periods such as holidays and vacations, making it important for users to understand pricing trends before booking. In this study, three different datasets were analyzed to extract meaningful insights into fare patterns, and the identified features were applied to seven different Machine Learning (ML) models to compare their predictive performance. The main objective was to identify the most influential factors and develop an effective prediction system. Among the applied techniques, Linear Regression, Decision Tree Regression, and Random Forest Regression were used as core models to enhance accuracy. Linear Regression provided a simple baseline by modeling linear relationships, Decision Tree Regression captured non-linear patterns through hierarchical data splitting, and Random Forest Regression improved robustness by combining multiple decision trees to reduce overfitting. By integrating these approaches, the system effectively captured both simple and complex relationships within the data, resulting in more reliable price predictions. This model can assist users in making informed booking decisions while also providing valuable insights into dynamic pricing behavior in the aviation domain.
Key words: Flight Ticket Price Prediction, Dynamic Pricing, Machine Learning (ML), Regression Models, Predictive Modeling, Data Analysis, Airline Pricing

1. INTRODUCTION

Air transportation has become a critical pillar of global connectivity, enabling efficient movement of individuals and businesses across geographical boundaries. Despite its widespread use, the pricing of flight tickets remains highly dynamic and unpredictable, posing challenges for travelers who seek cost-effective options. Ticket prices are influenced by a complex combination of factors, including demand variability, seasonal trends, route popularity, airline competition, booking lead time, and unforeseen external conditions such as holidays, economic shifts, or global disruptions. These continuously changing variables make it difficult for users to determine the optimal time to purchase tickets, often leading to uncertainty and suboptimal booking decisions.

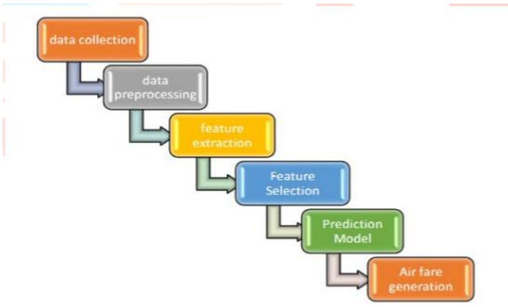

Fig. 1: Operational flow of Predicting Flight Ticket Fare

To address this issue, flight ticket price prediction has emerged as an advanced analytical solution powered by Machine Learning (ML) techniques. By leveraging historical pricing data, travel patterns, and feature relationships, ML models can identify hidden trends and forecast future fare movements with improved accuracy. These predictive systems analyze both linear and non-

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

linear dependencies among variables, enabling them to capture complex pricing behaviors that traditional statistical methods may overlook. Commonly used regression-based approaches, along with ensemble techniques, contribute to building robust models capable of handling large and diverse datasets.

The primary objective of such systems is to provide actionable insights that help travelers make informed decisions, such as identifying the most suitable booking window or selecting cost-efficient routes. In addition to benefiting consumers, these models also support airlines in demand forecasting, dynamic pricing strategies, and revenue optimization. Furthermore, practical implementation aspects such as scalability, data preprocessing, and model evaluation are essential to ensure reliable performance in real-world environments. Ethical considerations, including transparency, fairness, and avoidance of biased predictions, also play a significant role in the deployment of such systems. As ML-driven solutions continue to evolve, flight ticket price prediction holds the potential to transform the travel industry by enabling smarter decision-making, reducing uncertainty, and enhancing overall efficiency for both travelers and service providers.

Problem Definition:

The problem with flight ticket pricing lies in its volatility, influenced by numerous factors such as booking time, demand, seasonality, and Flight policies. This unpredictability makes it challenging for travelers to determine the optimal time to book tickets at the best price. The Flight Ticket Price Predictor aims to address this issue by using machine learning algorithms to accurately predict flight prices, helping users make informed and cost-effective booking decisions.

2. LITERATURE REVIEW

K. Tziridis, et al. [1] investigated the problem of flight ticket price prediction by utilizing multiple Machine Learning (ML) techniques and identifying key features that influence airfare. Their study selected various attributes such as departure details, route characteristics, and temporal factors to model pricing behavior. They experimented with several regression-based and neural network models and observed that feature selection played a crucial role in improving prediction accuracy. Their work demonstrated that ensemble-based approaches and regression models were effective in capturing pricing trends and improving forecasting performance. W. Groves, et al. [2] developed an intelligent agent-based system to optimize the timing of flight ticket purchases. Their approach employed Partial Least Squares regression along with feature engineering techniques such as lagged feature computation and model selection strategies. The study showed that incorporating temporal dependencies significantly improved prediction accuracy, and the proposed model effectively handled multicollinearity issues commonly found in airfare datasets. J. Santos Dominguez-Menchero, et al. [3] analyzed pricing behavior in airline markets and proposed a methodology to determine the optimal purchase time for flight tickets. They utilized non-parametric isotonic regression techniques to model fare variations and identify trends over time. Their results indicated that customers could minimize costs by delaying purchases within certain time windows, and the study quantified the economic impact of early and late booking decisions. S. Rajankar, et al. [4] conducted a comparative study on flight fare prediction using various Machine Learning algorithms. Their work involved preprocessing flight datasets and applying models such as K-Nearest Neighbors, Linear Regression, Support Vector Machine, Decision Tree, Gradient Boosting, and Random Forest. They evaluated model performance using metrics such as Mean Absolute Error, Mean Squared Error, and R-squared, and concluded that Decision Tree-based models provided superior performance in capturing non-linear pricing relationships. T. Wang, et al. [5] proposed a comprehensive framework for flight price prediction by integrating multiple datasets along with macroeconomic indicators. Their study applied advanced Machine Learning models such as Support Vector Machine and Extreme Gradient Boosting (XGBoost) to capture complex relationships between features. The results showed high predictive accuracy, with XGBoost achieving the lowest error rates and demonstrating strong generalization capability. T. Janssen, et al. [6] developed a predictive model based on Linear Quantile Mixed Regression to estimate optimal ticket purchasing time. Their study focused on analyzing pricing behavior for specific routes and incorporated features such as time before departure and travel day classification. The proposed approach effectively captured variability in price distributions and provided insights into optimal booking strategies. Wohlfarth, et al. [7] introduced a data mining-based approach for travel price forecasting by incorporating yield management principles. Their work utilized clustering and statistical techniques to analyze pricing trends and develop decision-support systems. The study emphasized the importance of understanding dynamic pricing mechanisms in the airline industry and provided tools to assist customers in making informed purchasing decisions.

V. Kimbhaune, et al. [8] developed a flight fare prediction system using Machine Learning models such as Random Forest, Decision Tree, and Linear Regression. Their objective was to predict optimal ticket purchase timing and improve decision-making for users. The models were evaluated using metrics such as Mean Absolute Error, Mean Squared Error, and Root Mean Squared Error. Although the system achieved moderate accuracy, the study highlighted that incorporating larger and real-time datasets could significantly enhance prediction performance. E. Bachis and C. A. Piga [9] analyzed pricing behavior in low-cost airline markets by investigating online price dispersion across digital booking platforms. Their study focused on how identical tickets are sold at varying prices depending on factors such as booking time, demand fluctuations, and consumer search patterns. The proposed analysis highlighted that airlines leverage real-time pricing strategies to maximize revenue and implicitly segment customers, demonstrating that online environments significantly increase price variability and market inefficiencies. P. P. Belobaba [10] introduced a foundational approach to airline revenue optimization through seat inventory control within yield management systems. His work focused on allocating seats across multiple fare classes while considering uncertain demand and the perishable nature of airline inventory. The proposed framework effectively balanced early low-fare bookings with late high-fare demand, providing a structured methodology that became the basis for modern revenue management systems in the airline industry. Y. Levin, J. McGill, and M. Nediak [11] developed a dynamic pricing model that incorporates strategic consumer behavior in competitive airline markets. Their study focused on understanding how customers adjust purchasing decisions by anticipating future price changes, especially in an oligopolistic environment. The proposed model demonstrated that accounting for consumer waiting behavior is critical for optimizing pricing strategies, as ignoring such behavior can significantly reduce expected revenue. B. Smith, J. Leimkuhler, R. Darrow, and Samuels [12] implemented a large-scale yield management system in a real-world airline environment to optimize revenue through advanced forecasting and optimization techniques. Their study focused on integrating demand prediction with booking control mechanisms to dynamically allocate seat inventory across fare classes. The proposed system significantly improved revenue performance, demonstrating the practical effectiveness of operations research techniques in commercial airline pricing and resource management.

3. PROPOSED SYSTEM

The data acquisition phase of a sophisticated flight price prediction system extends far beyond simple historical logs, often involving real-time web scraping and API integration to capture live market fluctuations. This stage transforms raw inputs into a high-dimensional feature space where temporal data is encoded using cyclical transformations—such as sine and cosine functions—to help the model understand the periodic nature of months and weeks. Beyond basic route details, advanced feature engineering incorporates macro-economic indicators like fuel price indices and currency exchange rates, as well as demand-side signals like search frequency and local holiday calendars, which collectively allow the system to account for sudden volatility that a standard linear model would overlook.

To achieve high predictive precision, the modeling stage typically moves past basic regression toward ensemble learning techniques like Gradient Boosted Decision Trees, such as XGBoost or LightGBM. These models are particularly effective at handling tabular data with complex, non-linear interactions between features, such as the relationship between a specific airline’s pricing strategy and the remaining seat capacity. During the training process, the system employs automated hyperparameter optimization and rigorous cross-validation to prevent overfitting, ensuring that the model remains robust across different seasons and routes. This ensemble approach allows the system to weigh the outputs of multiple algorithms, resulting in a "meta-prediction" that is significantly more reliable than any single model’s output.

The backend architecture is designed as a scalable microservices environment where the trained model is served through a low-latency API, enabling the frontend to provide near-instantaneous results to the user. This infrastructure is underpinned by a continuous MLOps pipeline that monitors for "data drift," a phenomenon where the model’s accuracy degrades because real-world pricing behaviors have shifted—due to events like a global pandemic or a major airline merger. When drift is detected, the system triggers an automated retraining workflow that ingests the most recent "ground truth" data, ensuring the predictive engine evolves alongside the market. This creates a self-healing system that maintains its relevance without constant manual intervention.

Finally, the recommendation component acts as a strategic intelligence layer that translates raw price forecasts into actionable consumer advice. By analyzing the variance and confidence intervals of its own predictions, the system can determine whether a current price is a statistical outlier or a standard fare, powering "buy" or "wait" notifications for the user. This layer often utilizes reinforcement learning to optimize the timing of these alerts, learning from historical "successful" saves to refine its future suggestions. The integration of this recommendation logic with a dynamic frontend ensures that the system is not just a passive forecasting tool, but a proactive financial advisor that helps users navigate the complexities of airline revenue management.

3.1 RANDOM FOREST

A random forest is an ensemble of decision trees as shown in Fig. 2. Each tree in the forest is trained on a random subset of the data and makes an independent prediction. The final prediction is based on the majority vote (classification) or the average (regression) of the predictions from individual trees. The Random Forest algorithm, when applied to Flight Ticket Price Prediction, works by creating a collection of decision trees to make predictions.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

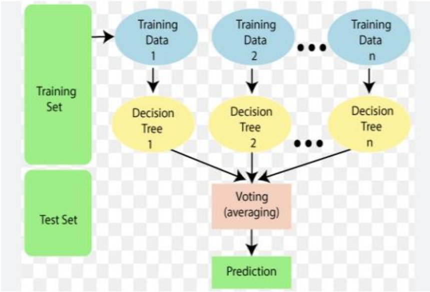

Fig. 2: Random Forest Structure

Each decision tree in the forest is built by randomly selecting subsets of features (such as flight route, booking time, seasonality, and travel class) and training on a random subset of the dataset. These trees work by splitting the data at various decision points based on feature values to predict the price, and each tree produces its own prediction. Once all the trees have made their predictions, the Random Forest algorithm aggregates the results, typically by averaging the predicted prices in the case of regression tasks, to generate a final output.

The key advantage of Random Forest is that it can handle complex, non-linear relationships between the features and the target variable (ticket price), unlike linear regression. Additionally, it is robust to overfitting due to the way it combines multiple trees and reduces variance. Random Forest can also handle high-dimensional data, dealing well with irrelevant features and missing data. This makes it a powerful tool for predicting Flight ticket prices, as it can capture intricate patterns in the data that simpler models might miss. However, the complexity of the model can make it less interpretable, and it may require more computational resources compared to simpler models.

4 CONCLUSION

In summary, the application of Machine Learning (ML) techniques to flight ticket price prediction has proven to be an effective approach for improving decision-making in the aviation domain. By utilizing historical pricing data and relevant influencing factors, predictive models are capable of identifying patterns and estimating future fare trends with reasonable accuracy. This enables travelers to plan their bookings more strategically and potentially reduce travel costs, while airlines can adopt data-driven pricing strategies to enhance revenue management and operational efficiency. The whole workflow involves key stages such as data collection, preprocessing, feature engineering, model training, and prediction, each contributing to the effectiveness of the system. Although challenges related to data inconsistency, feature selection, and model generalization remain, ongoing advancements in ML algorithms and the growing availability of large-scale datasets are expected to improve prediction reliability. Consequently, the integration of intelligent predictive systems into flight pricing mechanisms contributes to a more optimized, transparent, and user-centric travel ecosystem.

REFERENCES

[1] K. Tziridis, T. Kalampokas, G. Papakostas, and K. Diamantaras, “Flight ticket price prediction using machine learning techniques,” in Proc. European Signal Processing Conf. (EUSIPCO), 2017, doi: 10.23919/EUSIPCO.2017.8081365.

[2] W. Groves and M. Gini, “An agent for optimizing flight ticket purchasing,” in Proc. Int. Conf. Autonomous Agents and Multiagent Systems, 2013.

[3] J. Santos Dominguez-Menchero, J. Rivera, and E. Torres Manzanera, “Optimal purchase timing in the flight market,” 2014.

[4] S. Rajankar, N. Sakhrakar, and O. Rajankar, “Flight fare prediction using machine learning algorithms,” Int. J. Eng. Res. Technol. (IJERT), June 2019.

[5] T. Wang, S. Pouyanfar, H. Tian, and Y. Tao, “A framework for flight price prediction: A machine learning approach.”

[6] T. Janssen, “A linear quantile mixed regression model for prediction of flight

ticket prices," Bachelor Thesis, Radboud University, 2014.

[7] T. Wohlfarth, T. Clemencon, and S. Roueff, "A data mining approach to travel price forecasting," in Proc. Int. Conf. Machine Learning, Honolulu, 2011.

[8] Kalae, U. K. (2023). Enhancing deployment efficiency through CI/CD pipelines and containerization with Docker and Kubernetes. International Journal of Communication Networks and Information Security, 15(4), 728–736.

[9] Purmani, S. S. R. (2024). Aligning IT investment decisions with overall business strategy from an enterprise program management perspective, focusing on the integration of IT leadership in strategic decision-making processes. International Journal of Communication Networks and Information Security, 16(5), 1213–1219

[10] V. Kimbhaune, H. Donga, A. Trivedi, S. Mahajan, and V. Mahajan, "Flight fare prediction system using machine learning algorithms."

[11] E. Bachis and C. A. Piga. Low-cost Flights and online price dispersion. International Journal of Industrial Organization, In Press, Corrected Proof, 2011.

[12] P. P. Belobaba. Flight yield management. an overview of seat inventory control. Transportation Science, 21(2):63, 1987.

[13] Y. Levin, J. McGill, and M. Nediak. Dynamic pricing in the presence of strategic consumers and oligopolistic competition. Management Science, 55(1):32–46, 2009

[14] B. Smith, J. Leimkuhler, R. Darrow, and Samuels, ?Yield managementat american Flights,?Interfaces, vol.22, pp. 8–31, 1992.