Email: editor@ijerst.com or editor.ijerst@gmail.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 1, 2025

Dynamic pricing strategy for airline seat booking using SK learn

SK. Naga Rehmathunnisa[1], U. Sushma[2], CH. Maruthi Dhana Lakshmi[3], Sk. Rubeena[4], K. Satyavathi[5]

[1] Assistant Professor, Dept. of Computer Science & Engineering, Vijaya Institute of Technology for Women, Enikepadu, Vijayawada-521108

[2,3,4,5] Students, Dept. of Computer Science & Engineering, Vijaya Institute of Technology for Women, Enikepadu, Vijayawada-521108

Email id: shaikrehmathunnisa@gmail.com[1], usushma19@gmail.com[2], dhanalakshmi3149@gmail.com[3], shaikrubeena1702@gmail.com[4], rishitakureti@gmail.com[5],

Abstract:

This project aims to optimize airline seat booking through the implementation of a dynamic pricing strategy using machine learning techniques, specifically decision trees and random forests from the scikit-learn library in Python. The objective is to predict the optimal price for airline seats based on various factors such as demand, time until departure, seasonality, historical booking data, and competitor pricing. By leveraging these machine learning algorithms, the model can dynamically adjust seat prices in real-time to maximize revenue and occupancy rates. Decision trees provide insights into the factors influencing pricing decisions, while random forests enhance prediction accuracy by aggregating multiple decision trees. Through this approach, airlines can adapt their pricing strategies dynamically, leading to improved profitability and customer satisfaction.

Keywords: Decision trees,Scikit-learn, Random forests

1.Introduction

In today's fast-paced world, the airline industry stands as a testament to the complexities of modern commerce. Within this realm, the practice of pricing airline seats dynamically has emerged as a pivotal strategy for maximizing revenue and optimizing capacity utilization. Dynamic pricing, often hailed as the epitome of data-driven decision-making, leverages advanced algorithms and machine learning techniques to adaptively adjust seat prices based on a multitude of factors. This introduction delves into the intricate landscape of dynamic pricing strategies within the context of airline seat booking, with a specific focus on the utilization of the popular Python library, Scikit-learn (sklearn), for implementing machine learning models.

At the heart of dynamic pricing lies the fundamental principle of supply and demand dynamics. Airlines grapple with the challenge of efficiently managing their inventory of seats in response to fluctuating demand patterns, seasonal variations, competitor pricing strategies, and a myriad of other external influences. Traditional static pricing models, which offer fixed prices regardless of demand fluctuations, fall short in capturing the dynamic nature of consumer behavior and market conditions. In contrast, dynamic pricing empowers airlines to tailor their pricing strategies in real-time, thereby capitalizing on revenue opportunities while simultaneously ensuring a balance between supply and demand.

Sklearn, a prominent library in the Python ecosystem for machine learning, plays a pivotal role in enabling airlines to deploy sophisticated pricing models with relative ease and efficiency. With its rich suite of tools for data preprocessing, model training, and evaluation, sklearn provides a robust framework for building predictive models that underpin dynamic pricing strategies. Leveraging algorithms such as regression, classification, and ensemble methods, airlines can analyze historical

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

booking data, market trends, and customer segmentation to forecast demand and set optimal prices for airline seats.

2.Literature review

In the 1980s, American Airlines and its then president Robert Crandall started a revolution in airline pricing. Crandall is famous for many airline innovations in use today, such as devising the first frequent flier program and contributing to route optimization and central reservation system adoption. By offering cheaper rates to passengers who book earlier, providing seat reservation at a higher price, and incorporating the now controversial practice of overbooking, Crandall's AA claimed to generate an extra $500 million a year. The economic impact of yield management strategies utilized by American Airlines is apparent.

Today, revenue management specialists and systems work at most airlines to sell the right product to the right customer at the right moment at the right price on the right distribution channel. (Revenue management strategies in hotels are also common.) One of the critical components of revenue management is dynamic pricing. Let's talk about the current methods of dynamic pricing and how the techniques are maturing to match the current airline distribution market.

EXISTING SYSTEMS

Implementing dynamic pricing strategies for airline seat booking using machine learning techniques, particularly with the scikit-learn library in Python, involves a systematic integration process within the existing booking system. Beginning with data collection, historical booking data including flight details, pricing, and customer information is gathered and pre processed. Relevant features such as flight dates, times, route popularity, and competitor pricing are extracted through feature engineering. Subsequently, machine learning models, such as decision trees or random forests, are trained on this data to predict demand for airline seats. These models are then deployed within the booking system to provide real-time predictions based on current market conditions. Leveraging these predictions, seat prices are dynamically adjusted to optimize revenue and meet business objectives while ensuring customer satisfaction. Continuous monitoring and evaluation of the system's performance allow for refinement and adaptation to changing market dynamics, ultimately enhancing the airline's pricing strategies and profitability.

PROPOSED SYSTEM

This document outlines a system for predicting flight fares using machine learning techniques from scikit-learn, focusing on evaluating the prediction accuracy of Random Forest and Decision Tree algorithms.Airlines utilize dynamic pricing strategies to optimize revenue by adjusting ticket prices based on various factors. This system aims to predict flight fares with high accuracy using machine learning, enabling airlines to set competitive prices and maximize profits.The system requires historical flight data containing attributes that influence pricing. Publicly available datasets can be used for this purpose.The raw data needs cleaning and preparation for machine learning algorithms. This may involve handling missing values, encoding categorical variables, and feature scaling.This system will evaluate two popular algorithms from scikit-learn: Random Forest Regression and Decision Tree Regression. The models' performance is evaluated using metrics that measure how well they predict actual prices.Once evaluated, the model with the higher prediction accuracy will be used for prediction. Airlines can use this predicted price as a starting point for setting their final price, considering additional factors.

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

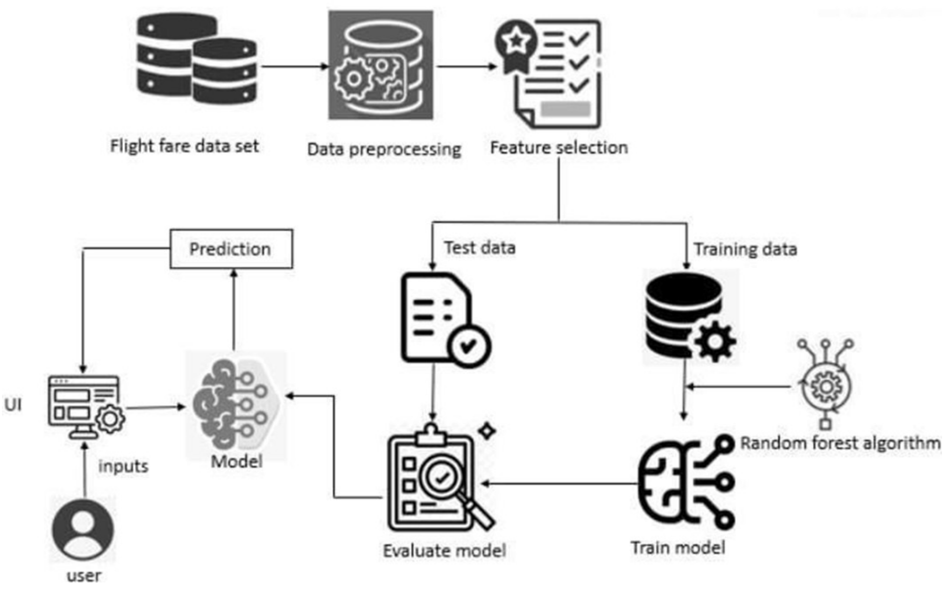
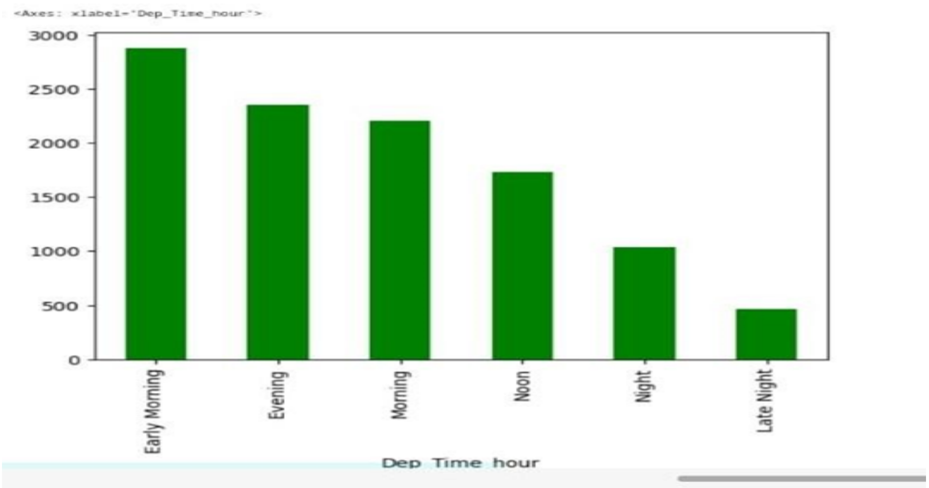
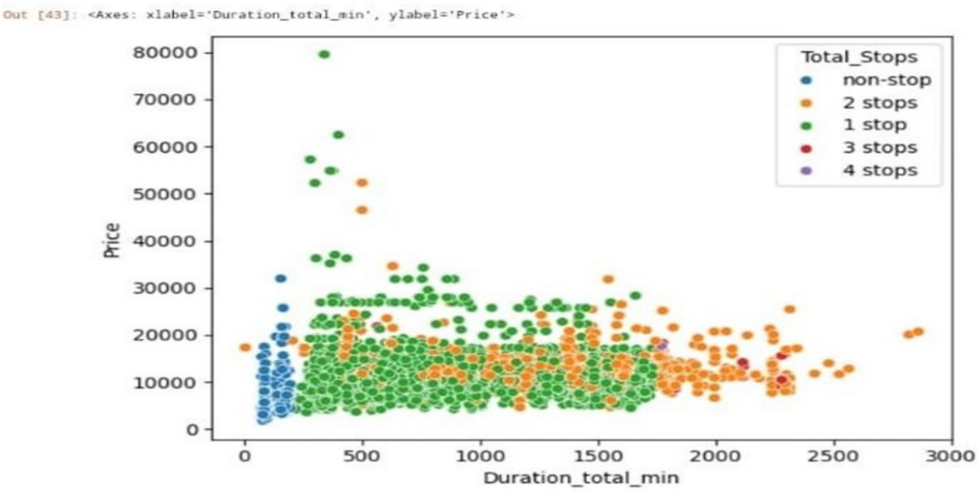
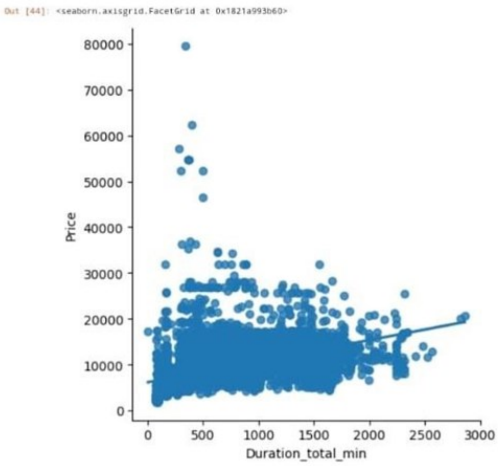


Figure: System Architecture





International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025



Future work:

Deep Learning Models: Explore the use of deep learning architectures such as recurrent neural networks (RNNs) or transformer-based models like BERT for improved prediction accuracy, especially in capturing complex temporal patterns and nonlinear relationships in the data.

Reinforcement Learning (RL): Investigate the application of reinforcement learning algorithms to enable the pricing system to learn and adapt its pricing strategy through interaction with the environment over time. RL can help optimize long-term revenue by dynamically adjusting prices based on feedback from past pricing decisions.

Dynamic Demand Forecasting:

Enhanced Demand Prediction Models: Develop more sophisticated demand forecasting models that incorporate a wider range of factors, including customer segmentation, market trends, and external events. Consider using techniques like time series analysis, ensemble learning, or Bayesian inference to improve forecasting accuracy.

Real-time Data Integration: Integrate real-time data streams from various sources such as social media, weather forecasts, and travel search engines to capture immediate changes in demand patterns and adjust pricing accordingly.

Data Processing and Analysis:

- Volume of Data: If dealing with large volumes of historical booking data, market demand data, and operational cost data, technologies like Apache Spark or distributed databases (e.g., Cassandra, Hadoop) may be necessary for efficient data processing and analysis.
- Variety of Data Sources: Utilize data integration platforms or tools like Apache NiFi or Talend to consolidate data from various sources such as airline reservation systems, market data providers, and external APIs.
- Velocity of Data: For handling real-time data streams and event-driven processing, technologies like Apache Kafka or AWS Kinesis can be used to ingest, process, and analyze streaming data in near real-time.

CONCLUSION

Dynamic flight fare prediction using decision trees and random forest methods, implemented with sklearn, offers a robust solution to the intricate pricing dynamics of the aviation industry. Flight fares are inherently dynamic, influenced by various factors such as demand fluctuations, seasonal patterns,

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

route popularity, and competitor pricing. Accurate fare prediction is essential for travelers seeking budget-friendly options, airlines aiming to optimize revenue, and travel agencies striving to offer competitive deals. Decision trees provide an intuitive framework for modeling complex decision-making processes by recursively splitting data based on features like departure date, time, destination, and airline. In contrast, random forests harness the collective intelligence of multiple decision trees to enhance predictive performance. By training on historical fare data and pertinent features, these methods produce accurate fare estimations. The versatility of decision trees and random forests enables them to adapt swiftly to changing demand patterns, evolving routes, and external influences, ensuring agility in pricing strategies. For travelers, dynamic fare prediction models facilitate informed decision-making by identifying optimal booking times, exploring alternative routes, and capitalizing on pricing trends. Airlines benefit from optimized revenue management strategies, maximizing profitability while ensuring competitiveness in dynamic markets. Moreover, these models empower airlines to forecast demand, allocate seats efficiently, and devise targeted promotions. Travel agencies can enhance customer satisfaction and loyalty by integrating dynamic fare prediction capabilities, providing real-time fare estimates, personalized recommendations, and insights into pricing trends. In conclusion, decision trees and random forest methods, implemented with sklearn, empower stakeholders in the aviation industry to make data-driven decisions, optimize pricing strategies, and enrich the travel experience for passengers globally.

References:

1. Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.
2. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media.
3. Müller, A. C., & Guido, S. (2016). Introduction to Machine Learning with Python: A Guide for Data Scientists. O'Reilly Media, Inc.
4. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of machine learning research, 12(Oct), 2825-2830.
5. Python Software Foundation. (2022). Python Language Reference, version 3.11. Available online: https://docs.python.org/3/