Modeling passenger preferences for air travel upgrades

Michel Robin, Bickerstaffe Faber and Magnin Antoine

University of Lausanne, HEC

Machine Learning in Business Analytics

2024

Group E

04.05.2024

Abstract: | The following machine learning project focuses on…

Introduction

- **Overview and Motivation**

In the fast-paced world of air travel, airlines face a dual challenge: maintaining operational efficiency while personalizing the travel experience for their passengers. Key aspects that greatly influence passenger satisfaction include baggage handling, preferred seating, and in-flight meals. Each of these elements represents significant opportunities for airlines to enhance customer experiences and optimize their ancillary services.

The project aims to harness machine learning to predict and analyze passengers' choices regarding baggage, preferred seating, and in-flight meal options. By examining a range of influencing factors, such as trip duration and purpose, the project seeks to generate actionable insights that airlines can use to improve operational efficiency and tailor their services to individual customer preferences.

Our motivation stems from the logistical and customer service challenges faced by airlines when handling passengers' preferences efficiently. Accurate prediction models will help airlines anticipate the demand for different services, from baggage needs to specific seating and meal preferences. This, in turn, will allow them to allocate resources more effectively and design better marketing strategies, enhancing customer satisfaction and maximizing revenue from ancillary services.

- **Data**

The dataset for our project is sourced from Kaggle and titled "Airlines Booking", compiled by user Anand Shaw. It is presented as a CSV file containing anonymized airline booking records, capturing a wide range of passenger data including flight details, baggage choices, preferred seat selection, and in-flight meal preferences. This comprehensive dataset provides the foundational data needed to analyze and understand the various preferences of air travelers, allowing us to identify patterns and predict future preferences. The dataset is accessible via Kaggle.

- **Related Work**

Previous research on passenger preferences in air travel has explored various factors that influence travelers' choices and behaviors. Key studies have investigated the impact of different service offerings, including baggage handling, seat selection, and in-flight meal preferences, on customer satisfaction and airline revenue.

Baggage Preferences:

Several studies have focused on the impact of baggage fees and allowances on customer behavior. Research by IATA (International Air Transport Association) and other organizations indicates that clear communication about baggage policies significantly influences booking decisions. The effect of ancillary baggage fees on travelers' booking choices and willingness to pay has also been explored, suggesting that transparency and flexibility in baggage policies can increase customer loyalty.

Preferred Seating:

Seat selection plays a critical role in enhancing the passenger experience. Studies have shown that passengers value proximity to exits, windows, or aisles, depending on their specific preferences. Research into the revenue impact of charging for preferred seating indicates that passengers are willing to pay extra for comfort and convenience, emphasizing the importance of predictive models that can help airlines cater to these preferences.

In-Flight Meals:

In-flight meal preferences have gained more attention as airlines seek to differentiate their services. Studies reveal that passengers have diverse dietary requirements and cultural preferences, influencing their satisfaction with airline services. Surveys and data analysis highlight the importance of offering a variety of meal options to meet these diverse needs, underscoring the value of predictive models that identify specific demands.

- **Research questions**

Our study, "Modeling Passenger Preferences for Air Travel Upgrades," focuses on developing predictive models to determine passenger choices for additional services during air travel. The central research question explores the application of machine learning:

How can machine learning models utilize passenger demographic and trip-specific data to predict preferences for air travel upgrades such as extra baggage, preferred seating, and in-flight meals?

This question aims to uncover the potential of using various data points to accurately forecast which upgrades passengers are most likely to select, thereby enhancing personalized service delivery and operational efficiency.

Data

- **Sources**

As previously introduced, our study utilizes the "Airlines Booking" dataset curated by Anand Shaw and hosted on Kaggle. This dataset, provided in CSV format, is essential for our analysis aimed at modeling passenger preferences for air travel upgrades.

- **Description**

`num_passengers`: Indicates the total number of passengers traveling on the booking.

`sales_channel`: Specifies the platform or method through which the booking was made.

`trip_type`: Describes the type of trip (e.g., Round Trip, One Way, Circle Trip).

`purchase_lead`: Represents the number of days between the booking date and the travel date.

`length_of_stay`: The number of days the passenger intends to stay at the destination.

`flight_hour`: The hour of the day when the flight is scheduled to depart.

`flight_day`: The day of the week on which the flight is scheduled.

`route`: The flight route from origin to destination.

`booking_origin`: The country from which the booking was made.

`wants_extra_baggage`: A binary indicator (yes/no) if the passenger opted for extra baggage.

`wants_preferred_seat`: A binary indicator (yes/no) if the passenger chose a preferred seating option during booking.

`wants_in_flight_meals`: A binary indicator (yes/no) if the passenger requested in-flight meals.

`flight_duration`: The total duration of the flight in hours.

`booking_complete`: A flag indicating whether the booking was completed (yes/no).

- Wrangling/cleaning
- Spotting mistakes and missing data (could be part of EDA too)
- Listing anomalies and outliers (could be part of EDA too)

```
# Example of a code block
```

Exploratory data analysis

- Mapping out the underlying structure
- Identifying the most important variables
- Univariate visualizations
- Multivariate visualizations
- Summary tables

Supervised learning

- Data splitting (if a training/test set split is enough for the global analysis, at least one CV or bootstrap must be used)
- Two or more models
- Two or more scores
- Tuning of one or more hyperparameters per model
- Interpretation of the model(s)

Unsupervised learning

- Clustering and/or dimension reduction

Conclusion

- Brief summary of the project
- Take home message
- Limitations
- Future work?