Name: Sri Varsha Naharajan
Semester: Fall 2024

Analyzing Customer Sentiments and Recommendation Factors in Airline Reviews(1416 words)

1. Introduction

The airline industry has long recognized the critical role of customer experience in fostering loyalty and driving recommendations. Whether it's the comfort of the seats, the friendliness of the staff, or the quality of the food, every detail plays a role in shaping how passengers feel about their journey. With the rise of online review platforms, travelers are now more vocal than ever about their experiences, sharing insights that airlines can use to improve their services. These reviews not only reflect customer satisfaction but also influence the likelihood of future travelers choosing a particular airline.

Researchers have long studied the factors that make passengers satisfied and loyal. For example, hang et al. (2023) explored how service quality—like responsiveness and reliability—affects customer satisfaction. Similarly, Bellizzi et al. (2020) highlighted key service quality factors in air travel, emphasizing the importance of passenger feedback. However, while many studies have analyzed structured ratings or survey responses, fewer have combined text-based sentiment analysis with numerical ratings to predict whether a customer would recommend an airline. This research fills that gap by not only predicting recommendations based on customer feedback but also uncovering recurring themes through topic modeling, providing airlines with actionable insights to elevate passenger experiences.

2. Research Question

How do different aspects of airline service, such as seat comfort, staff service, and in-flight entertainment, influence customer sentiments and their likelihood to recommend the airline?

3. Method

3.1 Data

The dataset for this study was sourced from the Airline Quality website (*Airline Ratings*, 2023) and includes 7,614 reviews from the top 10 airlines of 2023: Singapore Airlines, Qatar Airways, Emirates, Japan Airlines, ANA, Turkish Airlines, Air France, Cathay Pacific Airways, EVA Air, and Korean Air. The dataset contains 17 columns and was chosen for its combination of structured numerical ratings and unstructured textual reviews, offering a unique opportunity to understand customer sentiments and uncover the key factors that influence their likelihood to recommend an airline. The reviews were collected between early 2023 and early 2024, offering a recent and relevant perspective on customer sentiments.

Key features of the dataset include textual reviews written by customers, numerical ratings for critical aspects of service such as seat comfort, staff service, food and beverages, inflight entertainment, value for money, and overall satisfaction. Additionally, the dataset includes a binary "Recommended" variable, which indicates whether the customer would recommend the airline to others, serving as the target variable for supervised learning. Contextual information, such as the type of traveler (e.g., business or leisure), flight class (e.g., economy or business), route, and month flown, adds depth to the dataset, enabling a richer analysis of customer experiences across varying conditions.

3.2 Analysis

This study followed a structured approach, beginning with data cleaning and preparation, followed by sentiment analysis using Empath, recommendation prediction with a Random Forest model, and topic extraction using Latent Dirichlet Allocation (LDA). These methods effectively leveraged both numerical ratings and textual feedback to uncover key customer insights.

Data Preprocessing

Textual reviews were cleaned by removing special characters, punctuation, extra spaces, and stop words using the nltk.corpus module. Text was converted to lowercase for consistency, and non-alphanumeric characters were removed. Contextual information, such as traveler type, flight class, and route, was retained, while numerical ratings such as seat comfort, staff service, food and beverages, inflight entertainment, value for money, and overall rating were used as-is. The cleaned text was prepared for sentiment analysis and topic modeling.

Empath Sentiment Analysis

Sentiment analysis was conducted using the Empath library, which assigned normalized scores (0–1) to emotional categories such as positive_emotion, negative_emotion, trust, anger, contentment, fear, and joy. Scores above 0.5 in positive categories indicated satisfaction, while scores above 0.5 in negative categories signaled dissatisfaction. These sentiment scores were added to the dataset as numerical features, using pandas for data integration.

Supervised Learning with Random Forest

The Random Forest classifier, chosen for its ability to handle complex feature interactions and resist overfitting, was used to predict whether a customer would recommend the airline. The dataset was split into training (80%) and testing (20%) sets with a random seed (random_state=42) for reproducibility. The features used for this task included numerical ratings such as seat comfort, staff service, food and beverages, inflight entertainment, value for money, and overall rating, along with sentiment scores generated from Empath, including positive_emotion, negative_emotion, contentment, trust, anger, fear, and joy, while the target variable, *Recommended*, was encoded as a binary label (1 = yes, 0 = no). Grid Search optimized key hyperparameters, including n_estimators (50, 100, 200), max_depth (None, 10, 20), and min_samples_split (2, 5, 10). The scikit-learn library was used for implementation and tuning. Feature importance scores identified the most significant predictors of recommendations.

Unsupervised Learning with Topic Modeling

LDA was applied to the cleaned reviews to extract themes. The text was transformed into a document-term matrix using CountVectorizer, limiting the vocabulary to the 1,000 most frequent words. Five topics were extracted, with each topic represented by key words reflecting recurring themes like seat comfort, food quality, and staff service. The results were visualized using word clouds (wordcloud library) and bar charts (matplotlib library) to provide clear insights.

By combining data preprocessing, sentiment analysis, supervised learning, and topic modeling, this study provides a comprehensive and reproducible methodology for understanding customer sentiments and identifying key factors that influence airline recommendations.

4. Results

The Random Forest model, optimized with hyperparameters (max_depth=None, min_samples_leaf=1, min_samples_split=5, and n_estimators=200), demonstrated excellent predictive performance. The model achieved an accuracy of 94.07%, precision of 95.35%, recall of 93.35%, and an F1-score of 94.34%. The classification report (Table 1) indicates a well-balanced performance for both classes: customers who would recommend the airline (class 1) and those who would not (class 0).

Table 1: Random Forest Classification Report

Class	Precision	Recall	F1-Score	Support
0	93.00%	95.00%	94.00%	763
1	95.00%	93.00%	94.00%	857
Accuracy			94.07%	1620
Macro Avg	94.00%	94.00%	94.00%	1620
Weighted Avg	94.00%	94.00%	94.00%	1620

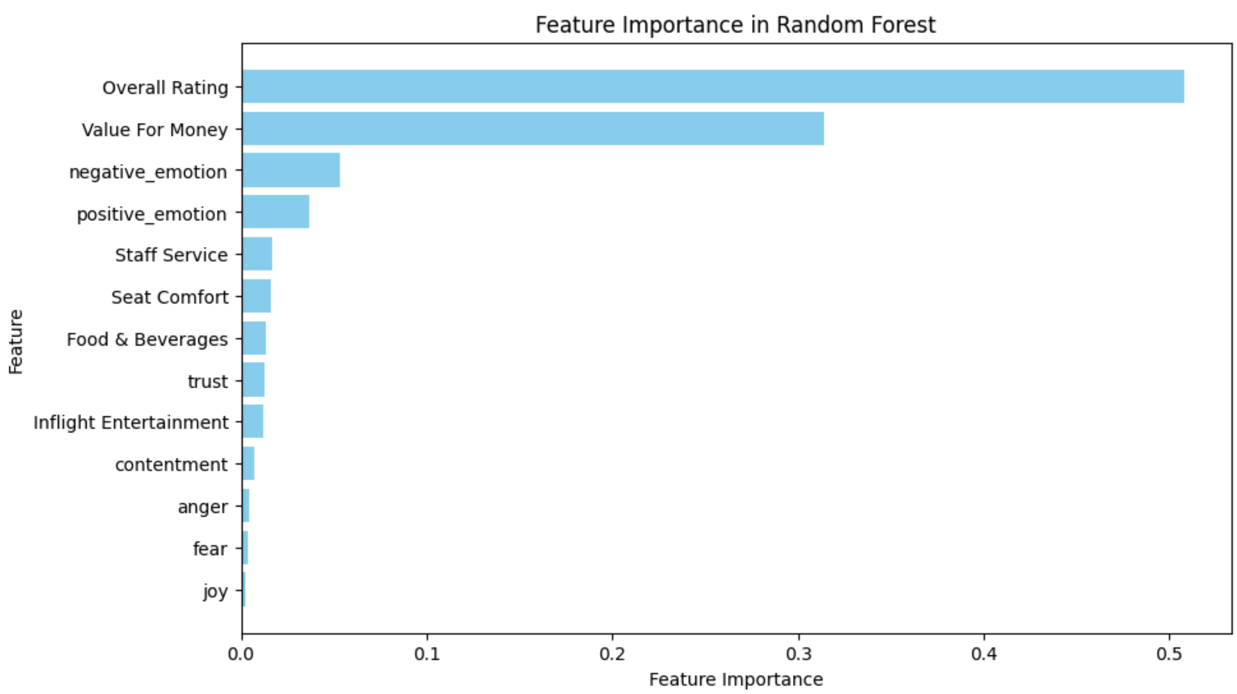
Feature importance analysis identified the most influential predictors of customer recommendations (Figure 2). Overall Rating was the most significant feature, contributing over 50% to the predictions, followed by Value For Money at 31%. Sentiment scores such as negative_emotion (5.3%) and positive_emotion (3.7%) also played a role, while features like trust, staff service, and food & beverages contributed smaller proportions. Table 2 includes detailed feature importance scores.

Table 2: Feature Importance Scores from the Random Forest Model

Feature	Importance
Overall Rating	0.5077
Value For Money	0.3137
negative_emorion	0.0530
positive_emotion	0.0366
Staff Service	0.0166
Seat Comfort	0.0164
Food & Beverages	0.0135
trust	0.0129

Inflight Entertainment	0.0121
contentment	0.0070
anger	0.0043
fear	0.0036
joy	0.0027

Figure 1: Feature Importance in Random Forest



The topic modeling analysis identified five key themes in the textual reviews, summarized in Table 3. These topics highlight important areas of customer feedback, such as premium experiences with Emirates (Topic 1), service quality and food with Qatar Airways (Topic 2), airport and delay issues (Topic 3), general feedback on food, service, and seating (Topic 4), and economy and premium class meals and service, particularly for Cathay Pacific and Singapore Airlines (Topic 5).

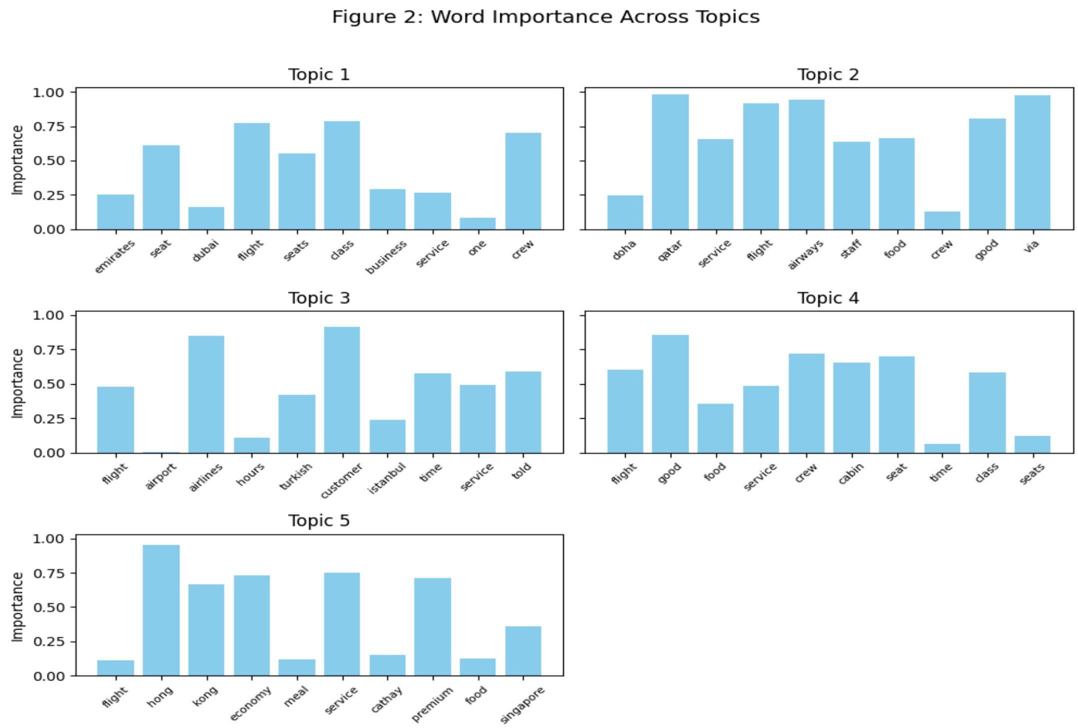
Table 3: Topic Modeling Results

Topic	Top Words
1	emirates, seat, dubai, flight, seats, class, business, service, one, crew
2	doha, qatar, service, flight, airways, staff, food, crew, good, via
3	flight, airport, airlines, hours, turkish, customer, istanbul, time, told
4	flight, good, food, service, crew, cabin, seat, time, class, seats
5	flight, hong, kong, economy, meal, service, cathay, premium, food, singapore

To provide a visual representation of the topics, word clouds (Figure 1) were generated to highlight the most frequent words in each topic, and bar charts (Figure 2) were used to display the importance of the top words, offering a clear understanding of the recurring themes in the reviews.

Figure 2 : Word Cloud for all the topics

Figure 3: Word Importance Across Topics



5. Conclusion and Limitations

The analysis revealed that overall rating and value for money were the strongest predictors of customer recommendations, highlighting the critical role of overall satisfaction and perceived affordability in driving passenger loyalty. Sentiment scores, particularly negative_emotion and positive_emotion, also contributed to the predictions, reflecting the influence of emotional responses on customer decisions. Positive sentiments like trust and joy were closely associated with favorable reviews, emphasizing passengers' appreciation for seamless experiences, professional staff, and premium services. In contrast, negative sentiments such as anger and dissatisfaction were linked to issues like delays, poor service, and discomfort. Features such as staff service, seat comfort, and food and beverages played smaller roles but remained relevant in shaping customer feedback. Topic modeling using Latent Dirichlet Allocation (LDA) provided qualitative insights into recurring themes, including premium travel experiences, food and service quality, and concerns about economy-class offerings and airport delays. These findings highlight the need for airlines to prioritize overall service quality and affordability while addressing operational inefficiencies to enhance customer satisfaction and increase recommendations.

Despite these valuable findings, the study has certain limitations. The dataset was limited to reviews from the top 10 airlines, which may not fully capture customer sentiments across the broader airline industry. Additionally, the sentiment analysis relied on predefined categories from the Empath library, which might not account for nuanced or airline-specific emotions. Finally, the analysis assumed all reviews were unbiased and authentic, while online platforms can sometimes reflect extreme opinions. Expanding the dataset and incorporating more advanced sentiment tools could address these limitations and further enhance the generalizability of the results in future research.

6. References

Airline Ratings. (2023). *Airline quality reviews*. Retrieved from https://www.airlinequality.com

Bellizzi, M. G., Eboli, L., & Mazzulla, G. (2020). Air transport service quality factors: A systematic literature review. *Transportation Research Procedia, 45*, 218–225. https://doi.org/10.1016/j.trpro.2020.03.010

Zhang, Y., Lee, S., & Gu, Y. (2023). A review of air transport service quality studies: Current status and future research agenda. *Journal of the Air Transport Research Society, 1*(1), 9–21. https://doi.org/10.59521/EF52BB6324BD7035