Study of Airline Passenger Preference Formation Mechanisms - An Application of a Hybrid Analysis Method Based on Skytrax Ratings and Reviews

Yichen Liu[1*] *Ye* Zhang[2]

[1]School of Finance, Guangdong University of Foreign Studies, Guangzhou, China
[2]School of Journalism and Communication, Anhui Normal University, Wuhu, China

Abstract. This study explores airline passenger preference formation by integrating structured rating data and unstructured review text from over 10,000 Skytrax user comments. A hybrid analytical framework combining binary Logit regression and natural language processing is employed to assess the impact of seven key service dimensions—including value for money, cabin service, and ground service—on passengers' willingness to recommend. Results show that value for money is the most influential factor, with notable differences in preference across traveler types and cabin classes. Sentiment analysis using a supervised learning model maps review emotions to a -5 to 5 scale, revealing a U-shaped distribution: extremely satisfied or dissatisfied passengers are more likely to leave reviews. Text mining further shows that dissatisfaction often relates to delays and pricing, while satisfaction is tied to crew service and comfort. The study offers actionable insights for airlines to identify service pain points, optimize experience for different passenger segments, and enhance customer satisfaction through data-driven decision-making.

1 Introduction

In recent years, driven by passengers' increasing awareness of service quality in the highly competitive and uncertain air transport market, a deep understanding of passenger preference mechanisms is crucial for airlines to improve service quality, optimize product design, conduct precision marketing, and ultimately enhance core competitiveness and profitability. In this context, the airport industry has had to change rapidly in response to travelers' accelerating demands and an increasingly competitive business environment [1]. Against the background of increasingly homogenized global airline services, consumers' subjective evaluation of airline service experience has become an important basis for influencing their brand preference and choice decision.

The massive impact of the COVID-19 pandemic has shown how heavily airports depend on the air transportation business and its two key customers for their survival [2]. In this current era, the growing popularity of online sites has made the voices of passengers louder

*Corresponding author: 20221704753@mail.gdufs.edu.cn

and more powerful, and their opinions are voiced more often to reflect their satisfaction regarding the service providers, including the airports themselves [3-5]. Airports have generally adopted the business model approach within their operation, regardless of ownership types [6].

Quality in service operations is a total experience usually evaluated by customers. Service operations always encompass multiple interactions between the customer and different employees. Therefore, the service industry must emphasize both differentiation and price [7]. The airline industry is very much influenced by changes taking place in its varied environment [8].

Satisfied passengers contribute to competitive advantages [9,10], share positive word of mouth, increase intention to reuse airports, boost non-aeronautical revenues, improve airports' reputation, and potentially influence travel plans to related destinations [11-13].

Amongst the famous online review sites for the airport industry are Google Reviews, the Skytrax website, and the individual airport's Facebook channels, where passengers can leave their comments and/or ratings of their airport experience after the journey. As a world-famous aviation service evaluation platform, Skytrax has gathered scores and comments from a large number of passengers on various service dimensions of airlines, making it an important data source for studying airline preference behavior. This study builds a multinomial logit model using Skytrax user comment data to evaluate and understand airline service quality. To identify and mitigate potential evaluation bias stemming from platform characteristics, this paper measures consumer airline preferences using evaluations across multiple service dimensions and analyzes the direct impact of these dimensions on preferences.

2 Experimental methodology

This study adopts a hybrid analysis method, combining a large amount of passenger rating data provided by the Skytrax website with textual comments, to conduct an in-depth investigation of how airline service preferences are formed. The data comes from the official website of Skytrax, covering more than 150 airlines around the world, with a total of more than 10,000 valid comment samples collected. By writing a crawler to extract the source code of the webpage and slicing the HTML structure, obtained key variables such as passenger ratings, comment text, passenger type, and airline company name. The data were cleaned and de-duplicated in several rounds to ensure the accuracy and representativeness of the analysis results.

In terms of rating modeling, the study used the traveler's willingness to recommend the airline he or she traveled with (Reco_bin) as the dependent variable and the seven dimensions defined by Skytrax as most representative of the airline's service (Seat Comfort(SC), Cabin Service(CSS), Food and Beverage(FB), Ground Service(GS), Value For Money(VFM), In-flight Entertainment(IE), and Wifi connectivity(WIFI))as the independent variables and used binary Logit regression analysis to assess the influence of each degree of influence of service dimensions on recommendation intention. To further identify the differences in service preferences among different groups of travelers, two moderating variables, class type (economy class, marked as 0, or business class, marked as 1) and traveler type (leisure travel, marked as 0, or business travel, marked as 1), are also introduced to construct an extended regression model containing the interaction term of the moderating term with the interaction term of the moderating term to examine the variability of the service dimensions in terms of service perception among specific groups.

In the text analysis section, the study will first extract review keywords using natural language processing techniques and TF-IDF methods combined with n-gram models, improve keyword recognition accuracy with the help of stemming and regular expressions, classify keywords, and evaluate the predictive power of keywords on recommendation

behaviors through logistic regression models. Based on the regression weights, the keyword dictionary is constructed, and the lexical categorization and weight visualization are carried out. Secondly, the sentiment analysis module will be introduced to build a sentiment scoring system based on the generated dictionary of keyword weights using logistic regression-based supervised learning techniques to map the sentiment from the least to the most satisfied to a satisfaction scale from -5 to 5 and to derive the link between the degree of recommendation and the degree of satisfaction to obtain the approximate distribution of travelers' sentiments, which will be evaluated by various correlation tests and model validation methods (e.g., AUC, accuracy, F1-score) to evaluate the effectiveness and generalization ability of the sentiment model. Finally, the keyword dictionary and sentiment distribution are jointly analyzed to identify the keywords or phrases that best reflect the positive/negative sentiments of travelers' comments and construct the dictionary. Based on the generated keyword sentiment polarity lexicon, the airline company can determine the aspects that are highly praised or in urgent need of improvement.

3 Empirical findings

3.1 Logistic regression analysis

Table 1. Descriptive statistical analysis of the seven dimensions (n=10,356)

Variables	Sample size (Obs)	Mean	Standard deviation	Minimum	Maximum
Reco_bin	10,356	0.272	0.445	0	1
SC	9,726	2.503	1.463	1	5
CSS	9,608	2.764	1.606	1	5
FB	7,239	2.414	1.495	1	5
GS	10,010	2.252	1.592	1	5
VFM	10,209	2.219	1.544	1	5
IE	5,846	2.382	1.506	1	5
WIFI	4,094	2.221	1.158	1	5

Table 1 presents a descriptive statistical analysis of the seven most critical dimensions measuring airline satisfaction on the Skytrax website. The results show that only about 27.2% of travelers indicated that they would like to recommend the airline they flew with, i.e., they have a preference for the airline company; most of the service dimensions' mean values are located between 2.2 and 2.7, which shows that travelers are satisfied with the individual services.

Table 2. Results of logistic regression without regulators (n=3,590)

Variables	Coefficients	Standard error	z-value	Significance
SC	0.429	0.095	4.54	0.000***
CSS	0.508	0.085	6	0.000***
FB	0.256	0.096	2.67	0.008***
GS	0.573	0.072	8	0.000***
VFM	1.229	0.097	12.62	0.000***
IE	0.267	0.081	3.29	0.001***
WIFI	0.006	0.008	0.67	0.501
Cons_	−10.857	0.471	−23.07	0.000***

Indicator	Value
LR chi² (df = 7)	3,138.36
Log Likelihood	−383.94
Pseudo R^2	0.8034

Table 2 shows the results of Logit regression on the data of the seven service dimensions, and it can be found that the explanatory power of the model is strong (Pseudo R^2=0.8034), indicating that the scoring variables can explain the variability of recommendation intentions better. Among all seven service dimensions, value for money (VFM) has the most significant positive effect and is the most central variable influencing passengers' recommendation behavior, followed by ground service (GS) and cabin service (CSS). In-flight entertainment and food, and beverage services also show some influence, while the effect of Wi-Fi services on recommendation intentions is not statistically significant.

After further introducing traveler type (Travel_bin) and cabin class (Seattype_bin) as moderating variables and constructing an extended model including interaction terms, the regression results are shown in Table 3. The explanatory power of the model is further improved (pseudo R^2 = 0.8109). However, none of the moderating variables were able to pass the confidence test at level 0.1, indicating that the two moderating variables were not significantly associated with traveler recommendation. The regression results of the interaction terms show that there is a significant moderating effect of passenger type on all evaluations related to cabin service: specifically, leisure travelers are significantly less sensitive to cabin service than business travelers but more responsive to food and beverage service. The more significant moderating effects of the class variable on WIFI and seat comfort suggest that business class travelers have higher expectations of the quality of the network and the comfort of the seat of the airline they are traveling with than economy class travelers, which may inhibit their willingness to recommend the service if it does not meet their expectations. These findings emphasize the heterogeneity of service perceptions across different groups of travelers and provide empirical evidence for precision marketing and service stratification by airlines.

Table 3. Logistic regression results with regulators and their interaction terms (n=3,590)

(1) Reco bin vs. Seven service dimensions		(2) Reco bin vs. Travel_bin interaction term		(3) Reco bin vs. Seattype_bin interaction term	
SC	0.330***	Travel_bin	-1.782	Seattype_bin	-2.965
	(0.006)		(0.241)		(0.104)
CSS	0.353***	Travel_SC	0.158	Seat_SC	0.290*
	(0.000)		(0.562)		(0.087)
FB	0.271	Travel_CSS	0.756***	Seat_CSS	0.304
	(0.022)		(0.007)		(0.363)
GS	0.697***	Travel_FB	-0.675**	Seat_FB	0.253
	(0.000)		(0.022)		(0.135)
VFM	1.178***	Travel_GS	-0.151	Seat_GS	-0.277
	(0.000)		(0.467)		(0.221)
IE	0.211	Travel_VFM	0.069	Seat_VFM	0.037
	(0.038)		(0.701)		(0.883)
WIFI	0.004**	Travel_IE	-0.209	Seat_IE	0.341
	(0.636)		(0.548)		(0.154)
Cons_	10.120***	Travel_WIFI	-0.299	Seat_WIFI	-0.340*
	(0.000)		(0.242)		(0.096)
Pseudo R^2	0.8109				
Log likelihood	-369.31336				
LR $chi^2(23)$	3167.61				

3.2 Comment keywords and sentiment analysis

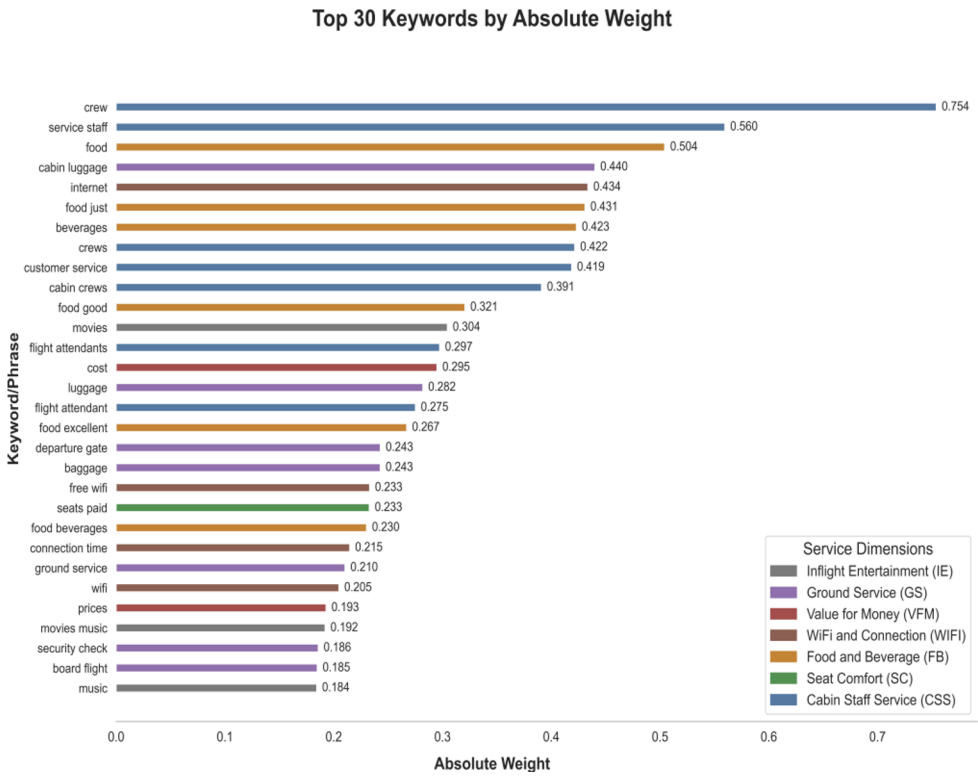

Fig. 1. Comment keyword weighting chart

Figure 1 demonstrates the results of the comment keyword weighting table. Due to the space limitation, this paper excerpts the 30 words with the highest weights in the travelers' comments for analysis. After assigning weights to the keywords through the model, it is found that "cabin crew" has the highest weight among all the keywords, which indicates that regardless of the type of travelers, the attitude and performance of the cabin crew are the most concerned and mentioned dimensions; other service dimensions that are mentioned more often include "food Other service dimensions mentioned more frequently include "food and beverages" and "ground service." It is worth noting that although value for money is the most influential factor in the ratings, it is seldom expressed directly in the comments. It is speculated that the reason may be that travelers have a vague perception of the abstract concept of "value for money" and tend to integrate it into their perception of the other six service dimensions, which ultimately leads to a relatively sparse direct discussion of value for money. The result is that direct discussion of value for money is rare. The discrepancy between travelers' concerns expressed in the text and the preferences presented in the ratings is revealed.

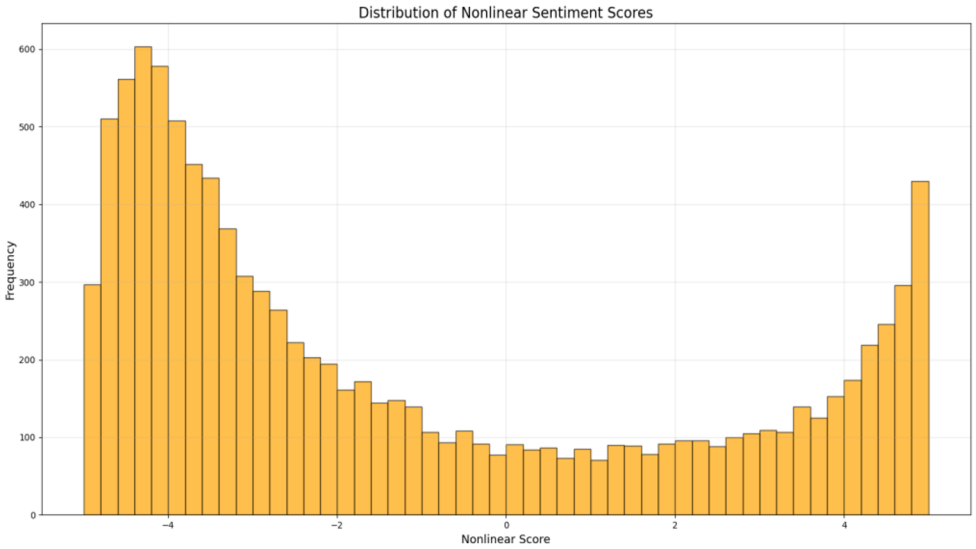

Fig. 2. Distribution of comment sentiment

Further, the study introduces a supervised learning model based on emotional polarity, and the resulting sentiment distribution is shown in Figure 2 above. After scoring travelers' emotions and mapping them to the [-5, 5] interval, it can be seen that the travelers' emotion scores show a clear U-shaped distribution, indicating that extremely satisfied and extremely dissatisfied travelers are more inclined to leave reviews, while emotionally neutral people have a weaker willingness to express their opinions. This polar characteristic reveals the pattern of user expression behavior driven by emotions and suggests that airline companies need to pay attention to the feedback of "disappointed" passengers in service management.

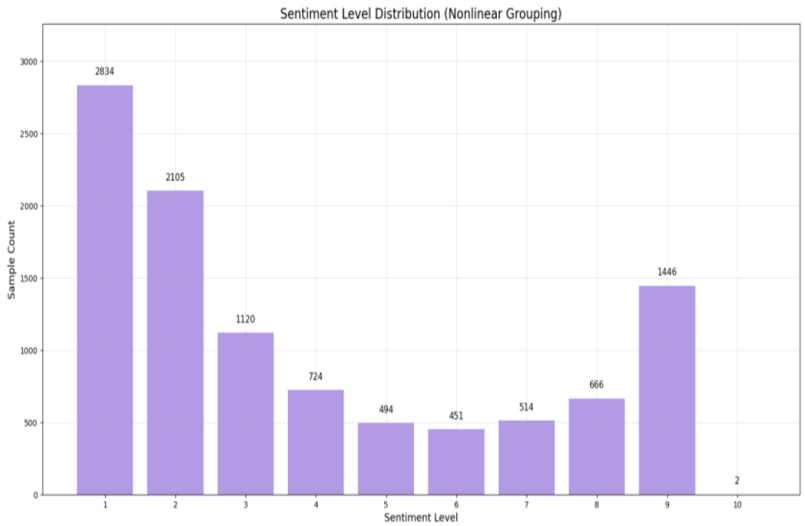

Fig. 3. Distribution of satisfaction scores mapped according to sentiment scores

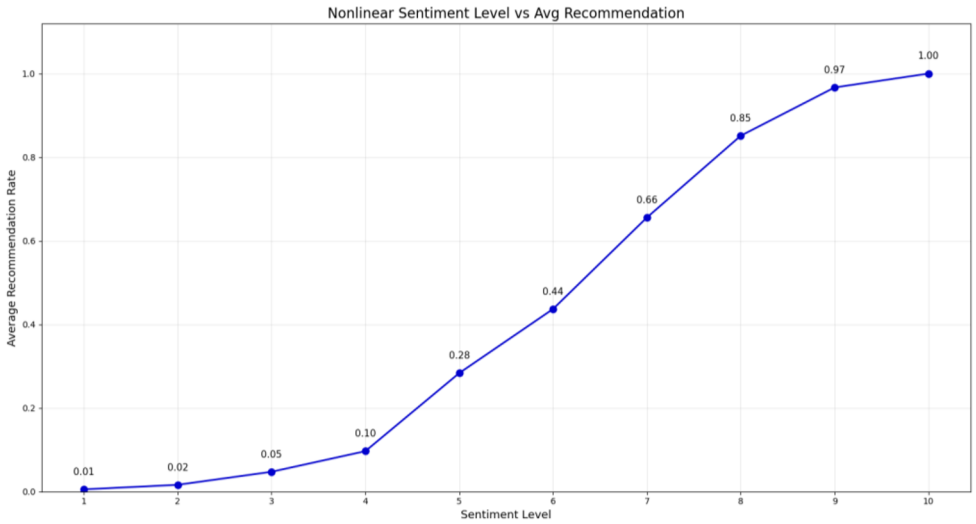

Fig. 4. Distribution of recommended expectations

Figures 3 and Figure 4 show the distribution of satisfaction scores and travelers' recommendation expectations based on the emotional scores. When the expectation value is greater than or equal to 0.5, it is believed that the travelers have a greater tendency to recommend the airline company at this time. According to the table, when the overall passenger satisfaction score reaches 7 or more, the expectation of recommendation is greater than 0.5, which can be regarded as a decision-making threshold mapping the satisfaction score to the "willingness to recommend," i.e., when the satisfaction score is greater than or equal to 7, the passenger will be inclined to recommend the airline. Since the Skytrax website will count the passengers' overall service satisfaction (10-point scale, with 1 being the least satisfied and 10 being the most satisfied) at the same time as it counts the degree of their recommendation of the airline, the analysis of the expectation of recommendation can not only help airlines to assess the degree of their passengers' preference when they have access to the data of satisfaction score, but also assist passengers to choose the airline with a high degree of recommendation when they only know the data of satisfaction score. It can also assist travelers in choosing the airline with a high recommendation level when they only know the satisfaction data. It is of great practical significance.

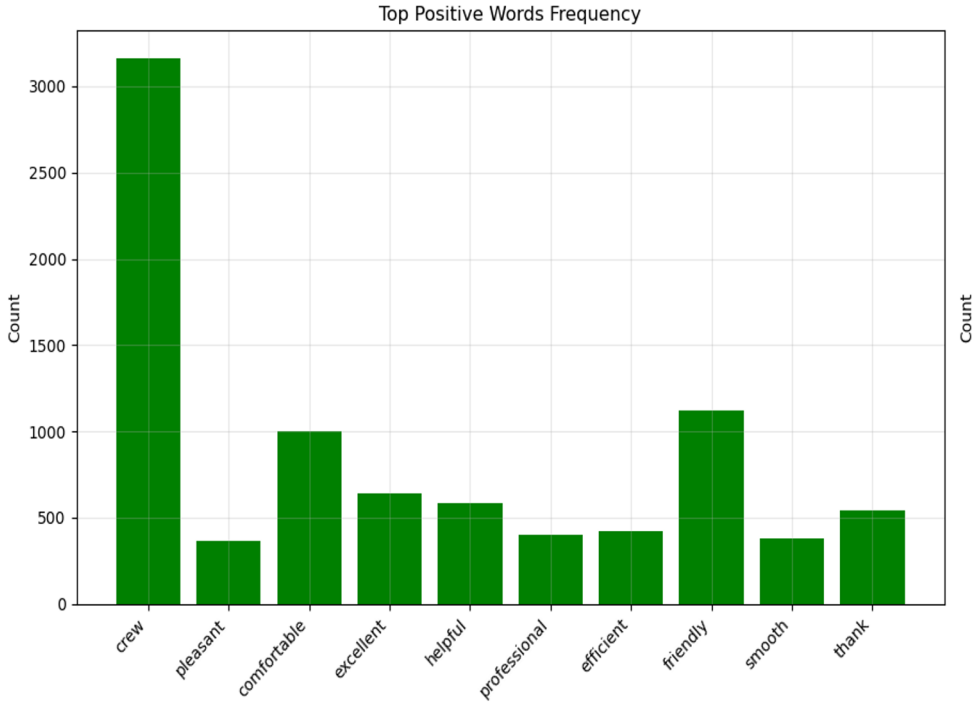

Fig. 5. Results of keywords containing emotional weights (Top Positive Words Frequency)

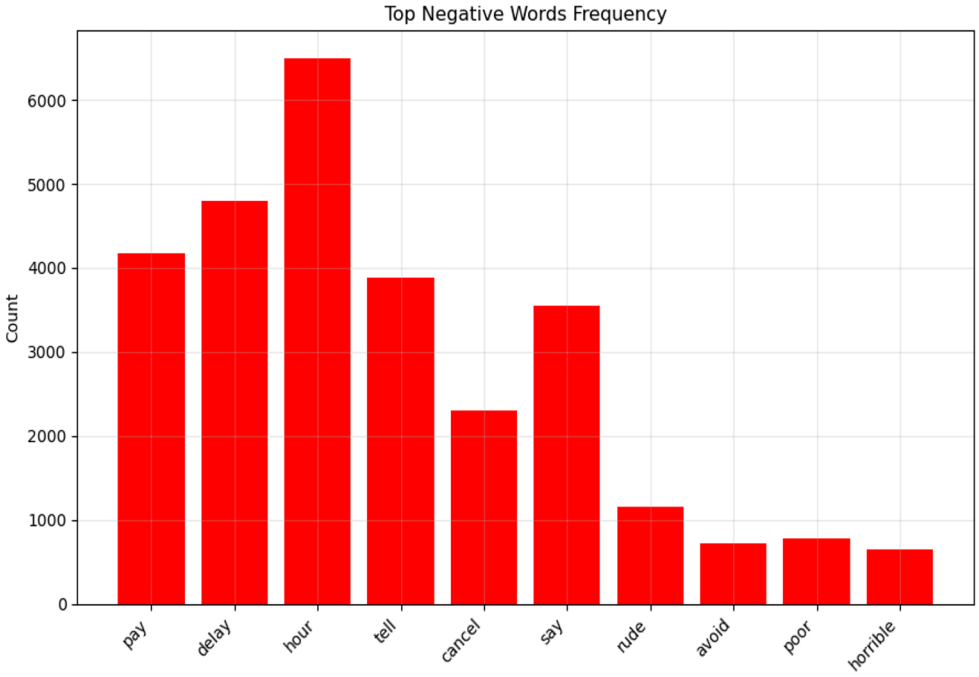

Fig. 6. Results of keywords containing emotional weights (Top Negative Words Frequency)

Figure 5 and Figure 6 show the results of integrating sentiment score modeling and keyword text mining. The grouping is based on the sentiment scores, with the negative-scoring samples labeled as "disappointed passengers" and the positive-scoring samples labeled as "satisfied passengers." The joint analysis of the two groups shows that satisfied travelers mainly focus on crew service and process (friendly, comfortable), while disappointed travelers focus on price (pay), delay (delay), and ground service (cancel). That is, words containing ground service types carry the disappointment of travelers. Considering the conclusion of the previous sentiment analysis that "passengers who are dissatisfied with the service are more likely to post comments than those who are satisfied with the service," believe that in the process of improving service quality, airline companies should prioritize the optimization of service nodes with a higher concentration of negative feedback from passengers and build a better negative emotion intervention mechanism. Therefore, believe that in the process of improving service quality, airline companies should prioritize the optimization of service nodes with a high concentration of negative feedback from passengers and build a better negative emotion intervention mechanism.

4 Conclusion

Based on the airline passenger rating and review data from the Skytrax platform, this study constructs a hybrid analysis framework integrating Logit regression modeling and textual sentiment analysis to systematically identify the service preference structure behind passengers' recommendation behavior. The empirical results show that cost-effectiveness is the most important core service dimension, which has the most significant positive influence on recommendation intention, followed by cabin service and ground service, which shows the key position of service experience in travelers' perception. By introducing passenger type (leisure/business) and cabin class (economy/business) as moderating variables, it is found that there is significant heterogeneity in the responses of different groups to the service dimensions: leisure travelers pay more attention to the quality of food and beverages, while business travelers are more concerned about the comfort of the seats and the stability of the network connection. In the text analysis section, the combination of TF-IDF and sentiment modeling identifies that the most frequently expressed points of dissatisfaction in passengers' comments are mainly focused on flight delays, the boarding process, and cancellation information, while positive sentiments are mostly focused on crew service attitude and ride comfort. The emotion scoring system indicates that the distribution of passenger emotions shows U-shaped characteristics, with negative emotions dominating the commenting behavior, and the probability of recommendation increases significantly when the satisfaction score reaches 7 points. Overall, this study jointly models passenger preferences from two paths: structured scores and unstructured text, which not only enhances the theoretical depth of airline service research in the dimension of emotion modeling but also provides an actionable empirical basis for airlines to accurately identify service pain points, optimize highly sensitive service links, and improve the effectiveness of customer relationship management.

Although this study is innovative in terms of data mining path and model integration techniques, there are still several limitations. First, in the process of emotion modeling, the study only adopts a simplified positive-negative binary grouping method to classify the comments into "satisfied" and "disappointed," which ignores the multi-dimensionality and delicacy of passengers' emotional expressions. Travelers often express not only simple satisfaction or dissatisfaction but also compound emotions such as anxiety, anger, surprise, appreciation, and so on, which are more complex emotional states that are not accurately distinguished in the current model, limiting the accuracy of emotion recognition. Second, the text analysis part mainly relies on TF-IDF and keyword frequency weights and lacks a

contextual semantic understanding mechanism, which makes it difficult to effectively capture semantic features such as metaphor, sarcasm, or double negatives in the comments, and there is a certain degree of risk of misjudgment. In addition, since the data analysis is still dominated by static review data and lacks dynamic tracking of travelers' behavioral processes (e.g., complaint follow-up, change in repurchase intention), the explanatory power in terms of behavioral prediction and feedback response mechanism is still limited.

Based on the above findings and shortcomings, future research can be further expanded in multiple directions. On the one hand, a more detailed multi-class emotion recognition system should be introduced in emotion modelings, such as the Plutchik roulette wheel model and the NRC multi-emotion lexicon, to enhance the ability to portray the heterogeneity of emotions; at the same time, pre-trained language models (e.g., BERT, RoBERTa, etc.) can be combined with deep learning to improve the model's comprehension of the complex textual structures and contexts and to further enhance the emotion assignment and further enhance the accuracy and interpretability of emotion assignment. On the other hand, it is recommended to integrate behavioral data, such as passenger complaint records, airline membership levels, repurchase behavior, etc., to build a comprehensive preference prediction model under the path of "emotion-cognition-behavior" to better support airline decision-making in customer management, crisis warning, and service iteration. This can better support airlines' decision-making practices in customer management, crisis warning, and service iteration. In addition, subsequent research can also explore the multi-platform integration mechanism by introducing user data from TripAdvisor, Google Reviews, and other platforms to validate the model's external applicability and cross-context stability and to promote the development of airline service satisfaction research from single-platform static evaluation to multi-source dynamic cognitive evolution.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

1. S.-J. Hong, D. Choi, J. Chae, Exploring different airport users' service quality satisfaction between service providers and air travelers. J. Retail. Consum. Serv. **54**, 101917 (2020)
2. K. Bunchongchit, Data analytics of Skytrax's airport review and ratings: Views of airport quality by passengers types. Procedia Comput. Sci. **181**, 563–570 (2021)
3. S. Gitto, P. Mancuso, Improving airport services using sentiment analysis of the websites. Tour. Manag. Perspect. **22**, 132–136 (2017)
4. K. Lee, C. Yu, Assessment of airport service quality: A complementary approach to measure perceived service quality based on Google reviews. J. Air Transp. Manag. **71**, 28–44 (2018)
5. B. Nghiêm-Phú, J.R. Suter, Airport image: An exploratory study of McCarran International Airport. J. Air Transp. Manag. **67**, 72–84 (2021)
6. A. Graham, Managing Airports: An International Perspective (5th edn., Routledge, London, 2020)
7. T.P. Van Dyke, L.A. Kappelman, Cautions on the use of the SERVQUAL measure to assess the quality of information systems services. Decis. Sci. **30**, 877–891 (1999)
8. E. Kaynak, O. Kucukemiroglu, A. Kara, Consumers' perception of airlines: A correspondence analysis approach in a global airlines industry. Manag. Int. Rev. **34**, 235–254 (1994)
9. D. Fodness, B. Murray, Passengers' expectations of airport service quality. J. Serv. Mark. **21**, 492–506 (2007)

10. W.-H. Tsai, W. Hsu, W.-C. Chou, A gap analysis model for improving airport service quality. Total Qual. Manag. Bus. Excell. **22**, 1025–1040 (2021)
11. G.C. Bezerra, C.F. Gomes, Antecedents and consequences of passenger satisfaction with the airport. J. Air Transp. Manag. **44**, 101766 (2020)
12. H. Barakat, R. Yeniterzi, L. Martín-Domingo, Applying deep learning models to twitter data to detect airport service quality. J. Air Transp. Manag. **91**, 102003 (2021)
13. S. Chatterjee, P. Kittur, P. Vishwakarma, A. Dey, What makes customers of airport lounges satisfied and more?: Impact of culture and travel class. J. Air Transp. Manag. **109**, 102480 (2023)