Satisfaction Analysis of Airline Passenger Experience

Yiran Zhang[a]
School of Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China

Keywords: Airline Satisfaction, XGBoost Model, SHAP Model, Visualization, Recommendations.

Abstract: With the rapid development of the aviation industry, airlines are increasingly focusing on passenger satisfaction with their flight experience. Not only does this enhance brand loyalty, it may also enable travelers to become effective evangelists for the brand. In order to deeply explore what affects air passenger satisfaction and the key factors behind it, this article studied a data set from Kaggle containing more than 120,000 airline passenger satisfaction scores. Through comparative analysis and verification, this paper selected the XGBoost model and SHAP model from many models. These two models have shown significant effects and accuracy in identifying and predicting factors related to overall satisfaction. At the same time, this article also uses a variety of data visualization methods to show the differences in satisfaction among different types of passengers and different routes. After detailed analysis of the visualization results and combined with market conditions, this article puts forward a series of targeted improvement suggestions, aiming to help airlines optimize services and improve satisfaction.

1 INTRODUCTION

With the development of the aviation industry and the process of globalization, competition in the aviation market has become increasingly fierce, and it is particularly important for airlines to manage passenger satisfaction. Aviation satisfaction will have a strong positive impact on repurchase and recommendation intentions (Mateja, 2017). By measuring passenger satisfaction, you can analyze the key factors that affect satisfaction, measure the current passenger satisfaction level, and make improvement strategies for products and services, thereby increasing the number of loyal customers and increasing market share and competitiveness. By building a satisfaction prediction model, passengers can be segmented into the market, different service products can be launched for different passengers, big data can be fully mined, and passenger satisfaction can be improved.

With the development of computer technology, many scholars use artificial intelligence methods such as machine learning to predict customer satisfaction. For example, through confirmatory factor analysis (CFA) and structural equation modeling (SEM), Chung found that airlines' corporate social responsibility activities can significantly affect passenger satisfaction, airline brand, and airline trust. In addition, passenger satisfaction and airline trust have a significant impact on airline loyalty (Chung, 2022). Spain's Mora and others have also revealed an assessment of customer satisfaction and loyalty levels of Colombia's low-cost and traditional airlines, arguing that the low-cost model reduces passenger satisfaction, thereby reducing loyalty (Garmendia，2021). At the same time, many articles also use the XGBoost model to analyze and predict satisfaction. Xu et al. integrated XGBoost and SMOTE algorithms to study which factors can be used to predict patient satisfaction (Xu, 2022). Gaofeng Guan et al apply the XGBoost model to rank the importance of each topic variable according to satisfaction (Guan, 2022).

Whether in medicine, supply chain or service industry, etc., it is not uncommon to use big data to analyze customer satisfaction (Sao, 2022; Alsayat, 2023).

The research route of this article is divided into four stages: data processing, model construction, visual analysis, conclusions and suggestions. This article intends to conduct research and analyze the overall satisfaction scores of 120,000 airline passengers, as well as the differences in the priorities

[a] https://orcid.org/0009-0002-6809-9425

Zhang, Y.
Satisfaction Analysis of Airline Passenger Experience.
DOI: 10.5220/0012911400004508

In *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2024)*, pages 141-152
ISBN: 978-989-758-713-9

and personalized needs of different groups of passengers for air travel. On this basis, conclusions are drawn and recommendations are made.

2 DATA PROCESSING

First, the article will conduct a simple summary analysis of the data. In order to facilitate data processing, the author named the 14 services as categories 1-14, as shown in table 1:

Table 1: Data column name.

Class	Name
1	Departure and Arrival Time Convenience
2	Ease of Online Booking
3	Check-in Service
4	Online Boarding
5	Gate Location
6	On-board Service
7	Seat Comfort
8	Leg Room Service
9	Cleanliness
10	Food and Drink
11	In-flight Service
12	In-flight Wifi Service
13	In-flight Entertainment
14	Baggage Handling

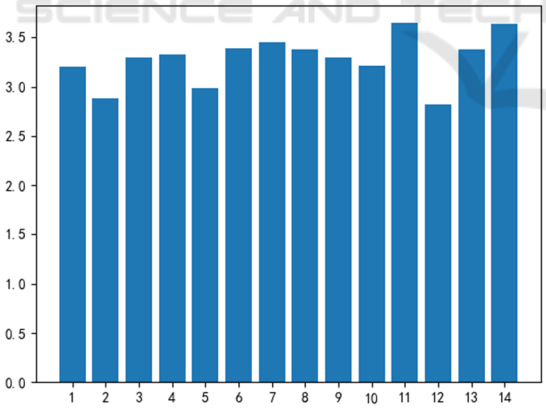

Figure 1a: Average statistics.

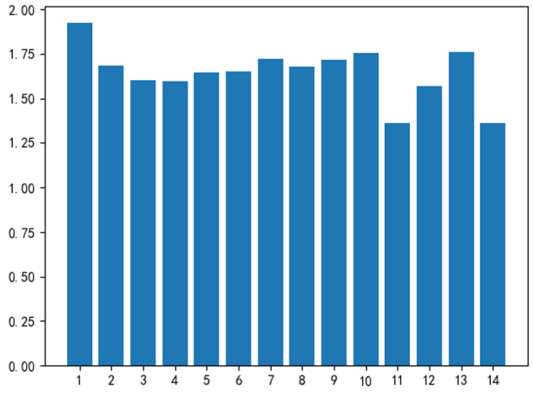

Figure 1b:Variance Statistics.

The analysis found that in-flight service and baggage handling had higher average ratings, while the convenience of online reservations, gate location, and in-flight Wi-Fi service had lower average ratings, as shown in Figure 1a and Figure 1b. Wifi service had lower average ratings. Meanwhile, departure and arrival times have the largest variance for convenience, while in-flight service and baggage handling have the smallest variance.

At the same time, the overall satisfaction distribution chart is shown in Figure 2:

The author found that the overall satisfaction level is neutral or there are more dissatisfied passengers than satisfied passengers, indicating that the overall satisfaction level is not high and there are many problems to be solved.

The correlation coefficient between different types of services is small and the correlation is weak. Therefore, different types of services can be regarded as independent of each other, and only the impact of each service on the overall satisfaction level is considered, as shown in Figure 3.

3 MODELING ANALYSIS

3.1 Feature Correlation Analysis

In order to study the impact of each service on overall satisfaction, the author calculated the correlation coefficient between them and drew a heat map as Figure 4:

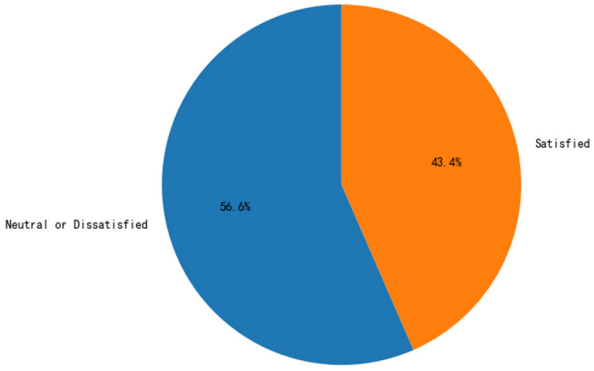

Figure 2: Overall satisfaction distribution chart.

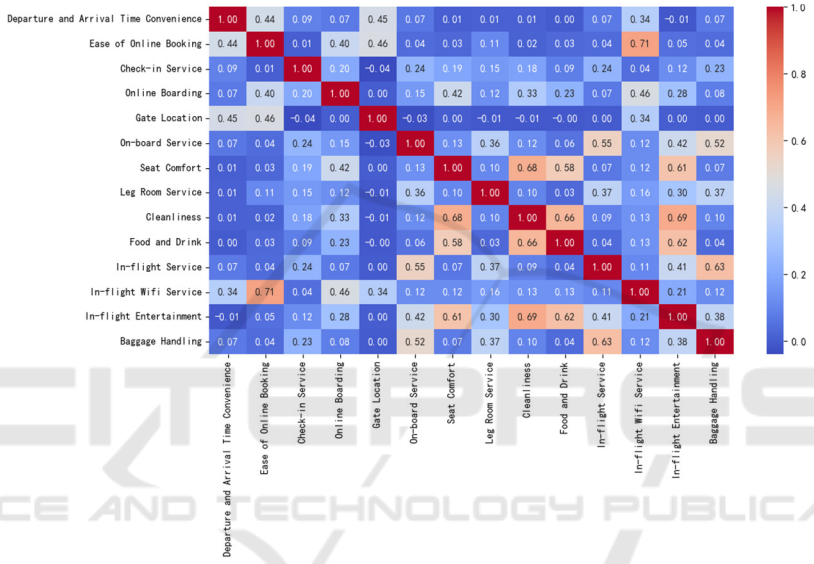

Figure 3: Heat map of correlation coefficients of various types of services.

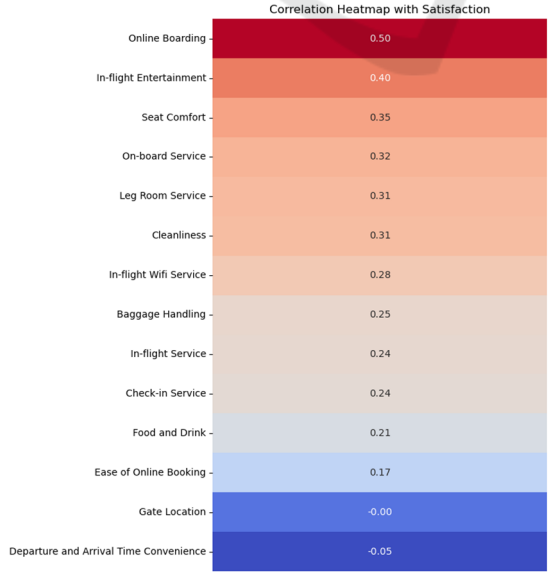

Figure 4: Heat map of the correlation coefficient between each category of services and overall satisfaction.

It is easy to find from the chart that online check-in, in-flight services, and comfort have higher correlation coefficients and have a greater impact on overall satisfaction. The convenience of online reservations, boarding gate location, and departure and arrival times have the lowest correlation coefficients and have the smallest impact on overall satisfaction.

3.2 Model Selection

In order to select the appropriate model, the author uses five different models: Logistic Regression, Random Forest, Adaptive Boost, Categorical Bayes, and XGBoost to evaluate and compare the accuracy of each service. The ratio of the number of test sets and training sets is 1 :4, the Precision and Recall rates of each model are in table 2.

The precision and recall rates of each model are visualized as shown in the Figure 5.

Obviously, each value of the XGBoost model is the best among the five models. Therefore, it is judged that the XGBoost model meets the situation of airline passenger satisfaction and is the most suitable for processing the data set of this question.

3.3 XGBoost Importance Analysis

Using the XGBoost model to analyze the importance of each service, the results are shown in the Figure 5.

Table 2: Comparison of model performance indicators.

	Model	Accuracy on Training	Accuracy on Test	ROC AUC Score	Precision	Recall
1	Logistic Regression	0.877097	0.877101	0.944868	0.860610	0.846040
2	Random Forest	0.929014	0.930794	0.977088	0.924671	0.910297
3	Adaptive Boost	0.907251	0.906715	0.963083	0.901406	0.876238
4	Categorical Bayes	0.893242	0.893093	0.949520	0.874839	0.871980
5	XGBoost	0.942293	0.951376	0.986617	0.943811	0.918020

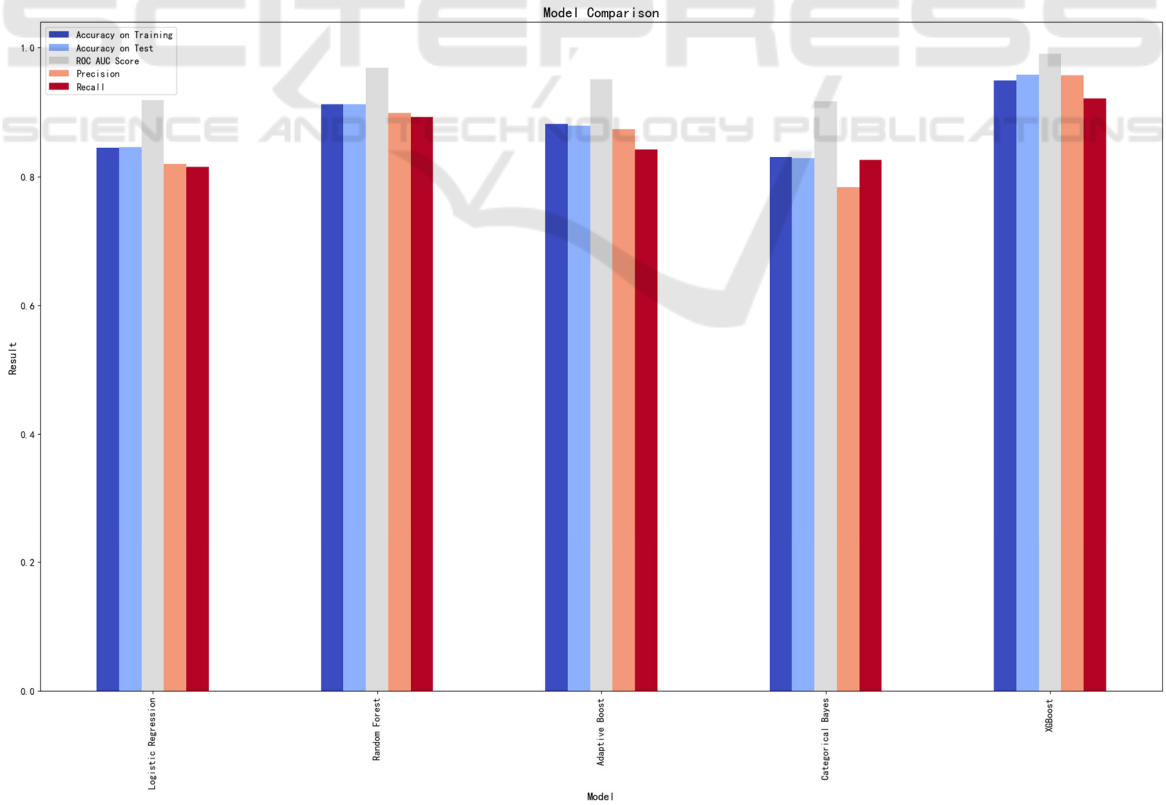

Figure 5: Model performance index comparison chart.

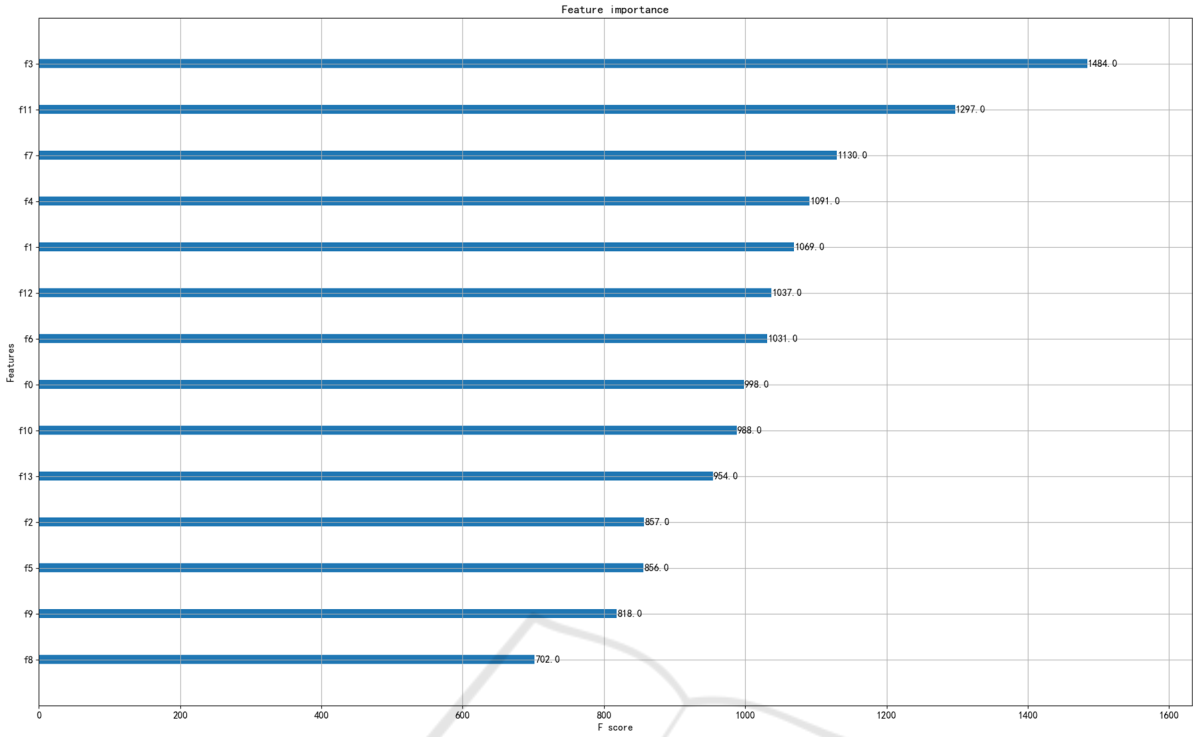

Figure 6: XGBoost model running results.

In Figure 6, it can be found that online check-in and in-flight Wi-Fi service are the most important, while food and beverages and cleanliness have the least impact on overall satisfaction. Service categories with high importance should be focused on improving, while service categories with less impact can be appropriately ignored.

3.4 SHAP Model Prediction

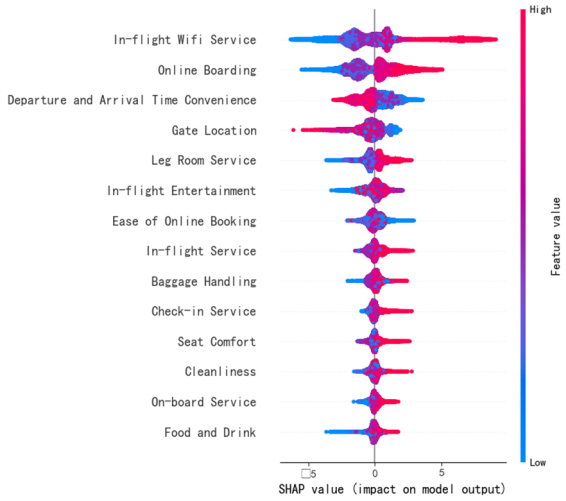

Figure 7: SHAP model running results.

Using the average SHAP value as an indicator of feature importance, as shown in Figure 7. Hyunwoo Park and Hyunju Jeon used the SHAP model to identify characteristics of customer satisfaction dimensions (Hyunwoo, 2022). It can be seen that the in-flight Wifi service has the farthest radiation range around zero point, indicating that it is the most influential feature. Most of the points for items such as food and beverages are piled at the zero point, indicating that their importance is the lowest.

3.5 Summary

Through the construction and operation of XGBoost and SHAP models, the author obtained the following results: among various services, the importance of online check-in, leg space service and in-flight Wifi service are all at the forefront. Therefore, it is recommended that airlines can reduce queuing time and improve passenger satisfaction by providing a more convenient online check-in process, including advance seat selection, online check-in, etc.; provide wider seat space to provide more comfortable To improve the riding experience, you can also consider adding adjustable seats and leg support facilities to meet the needs of different passengers; provide stable and high-speed in-flight wireless network connections to meet the needs of passengers, and

regularly maintain and upgrade in-flight wireless network equipment. to ensure proper functioning and good coverage.

Through the above measures, higher priority service categories can be improved, thereby increasing overall satisfaction.

4 VISUAL ANALYSIS OF DIFFERENT TYPES OF PASSENGER

Sudhakar once distinguished between passenger types and airline types to analyze customer satisfaction (Sudhakar, 2020). This article will analyze the satisfaction data of different types of passengers in five aspects to determine which types of passengers need to improve which service levels.

4.1 Analysis of Differences in Satisfaction Among Passengers of Different Ages

First, the data set of more than 120,000 passengers was cleaned to remove data with a score of "0" in various services. Then, a statistical chart is constructed with the age of passengers as the horizontal axis and the proportion of satisfaction and neutral or dissatisfaction as the vertical axis, as shown in the Figure 8.

In order to further analyze the differences in satisfaction of passengers of different ages for different categories of services, the average scores of various services for each age were calculated and visualized in three dimensions. At the same time, the average scores of passengers for different categories of services were calculated for different age groups based on the difference in overall satisfaction. As shown in the Figure 9 and Figure 10.

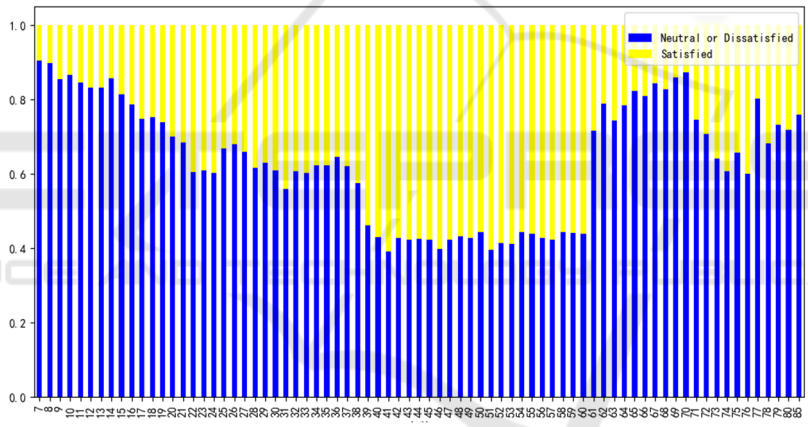

Figure 8: Proportion chart of overall satisfaction among travelers of different ages.

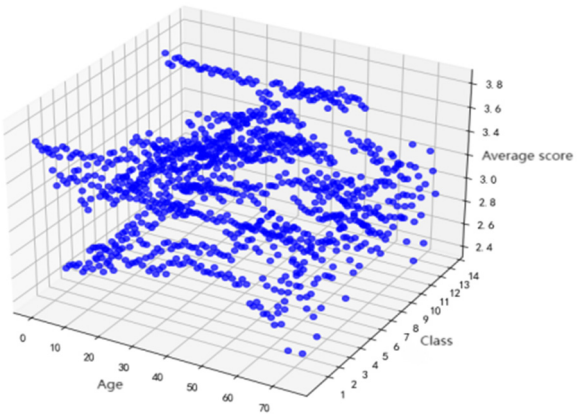

Figure 9: Scatter plot of average ratings of different services by travelers of different ages.

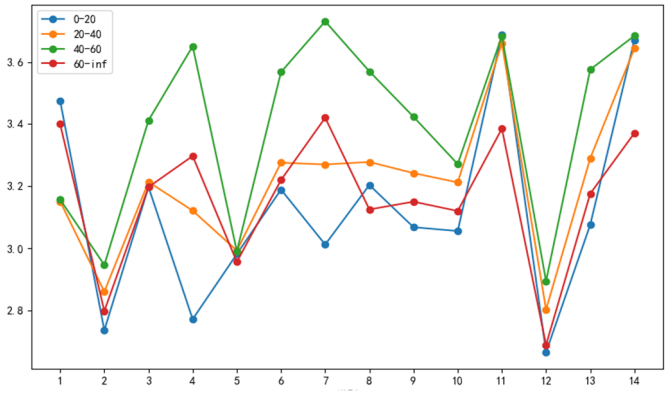

Figure 10: Line chart of average ratings of different services by travelers of different age group.

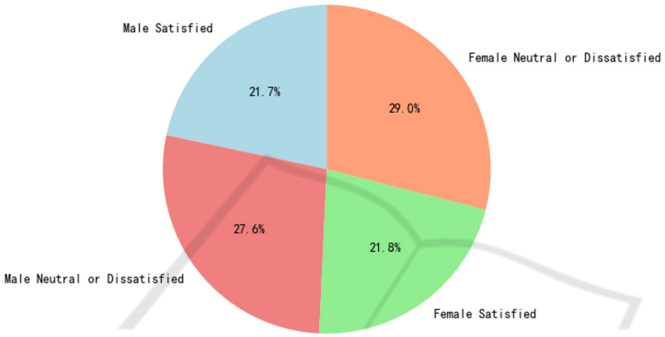

Figure 11: Distribution chart of overall satisfaction of passengers by gender.

It can be found in the Figure that the satisfaction and average ratings of people in the 40-60 age group are higher than those of other age groups in each rating item, while the satisfaction and average ratings of people in the 0-20 age group are generally the lowest. In addition, people over 60 years old have lower ratings for in-flight services, in-flight entertainment and WIFI services.

The reason for the lower ratings of passengers in the 0-20 age group may be a lack of travel experience, not accustomed to aviation services, or low tolerance, and the rating standards are relatively stricter. At the same time, for older passengers, there may be certain difficulties in using new technological equipment and services, and in-flight entertainment items may have certain requirements for vision or hearing, and older people may not be interested in the entertainment content provided on board. For example, the movies and TV programs provided by airlines may not meet the preferences of passengers over 60 years old, resulting in lower ratings.

It is recommended that airlines should understand the needs and preferences of passengers of different age groups and provide personalized services. For example, provide more entertainment facilities and Internet services for young passengers, and provide more comfort facilities and auxiliary services for elderly passengers; Providing more fast food and snack options for younger travelers and healthier and more digestible dining options for older travelers; Provide entertainment facilities such as movies, music and games for young passengers, and provide reading materials and relaxation space for elderly passengers. Airlines can use the above measures to improve the satisfaction of passengers of different age groups and provide a better travel experience.

4.2 Analysis of Differences in Satisfaction among Passengers of Different Genders

Data is extracted based on three elements: gender, overall satisfaction, and average rating of each service, as shown in the Figure 11 and Figure 12.

As can be seen from the Figure 11 and Figure 12, the proportion of male and female passengers who express satisfaction is almost equal to the total number. The proportion of female passengers who express neutrality or dissatisfaction is slightly higher than that of males, but there is no significant difference between the two. And in the average ratings of various services, gender did not cause

significant differences in ratings. It can be determined that gender has no significant impact on overall satisfaction and ratings.

4.3 Analysis of Differences in Satisfaction among Travelers of Different Customer Types

As shown in the Figure 12 and Figure 13.

As you can see from Figure 13 and Figure 14, the number of return travelers is approximately four times the number of first-time travelers. Among first-time travelers, the number of neutral or dissatisfied travelers is about three times the number of satisfied travelers, while among returning travelers, the number of neutral or dissatisfied travelers is roughly equal to the number of satisfied travelers. Among the average ratings of various services, except for the three categories of boarding gate location, baggage handling and in-flight services, the average ratings of returning passengers are slightly lower than those of first-time passengers. The average ratings of other returning passengers are significantly higher than those of first-time passengers. Among them, the average ratings of returning passengers are significantly higher than those of first-time passengers. The difference between space services

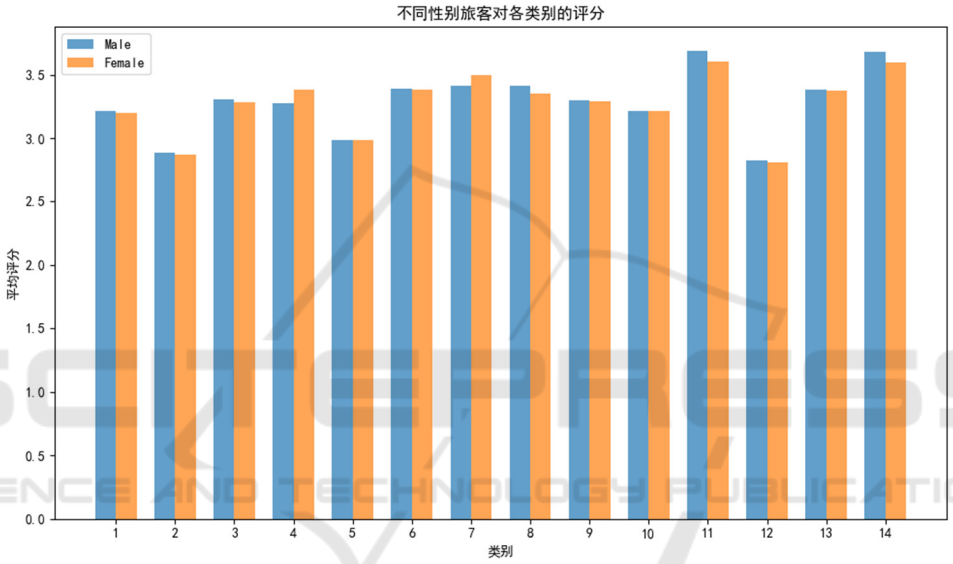

Figure 12: Histogram of average ratings of passengers of different genders.

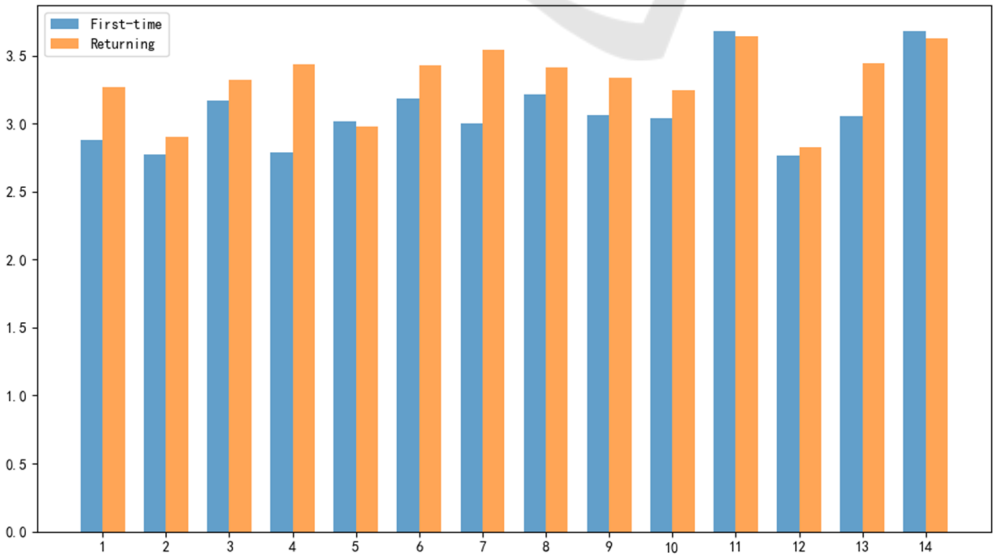

Figure 13: Histogram of average ratings of travelers by different customer types.

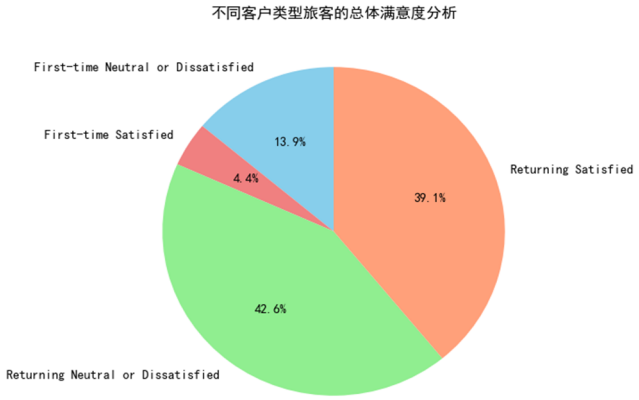

Figure 14: Distribution chart of overall satisfaction of passengers by different customer types.

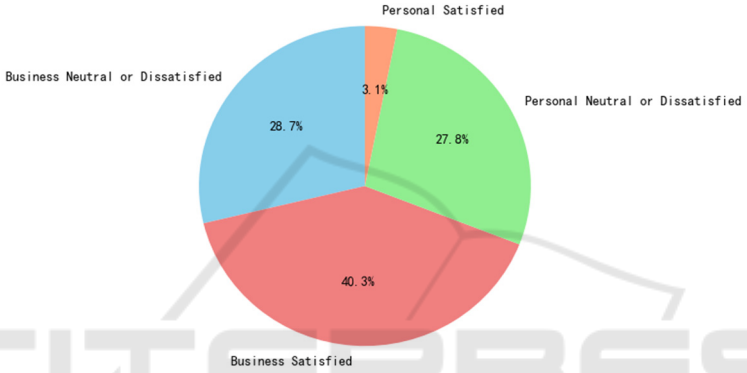

Figure 15: Distribution chart of overall passenger satisfaction for different flight destinations.

and online check-in is most obvious.

The reason for the above differences may be that passengers who fly with this airline again may have more travel experience. They have clearer understanding and expectations of the airline's aviation services. Compared with first-time passengers, they can be more familiar with the relevant service procedures and better evaluate and compare the service quality of different airlines.

Therefore, it is recommended that airlines should focus on and solve the needs of passengers experiencing their own flights for the first time and provide more information. For example, during the booking process, provide first-time users with clear information about the legroom size and configuration of each seat to help them make a better choice; Provide detailed boarding instructions on the airline's website or app, including times, locations and steps, to help first-time users understand the entire boarding process.

4.4 Analysis of Differences in Satisfaction Among Passengers with Different Flight Destinations

As shown in the Figure 15 and Figure 16.

The number of business travelers is about twice that of individual travelers, and the number of satisfied business travelers is significantly higher than the number of neutral or dissatisfied business travelers. Among individual travelers, the number who were neutral or dissatisfied was about nine times the number who were satisfied. Except for the convenience of departure and arrival times and check-in services, the average ratings of business travelers on other types of services are significantly higher than those of individual travelers.

The above differences may be due to the fact that airlines generally invest more resources and efforts in providing services to business passengers, including more spacious and comfortable seats, additional benefits (such as priority boarding, VIP lounge access) and higher-level dining and entertainment facilities, etc. After business travelers receive better services, they pay more attention to items that waste

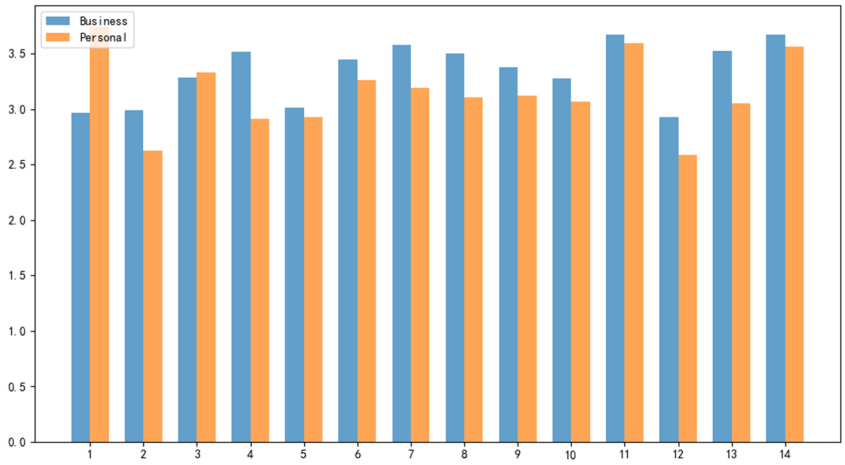

Figure 16: Histogram of average passenger ratings for different flight destinations.

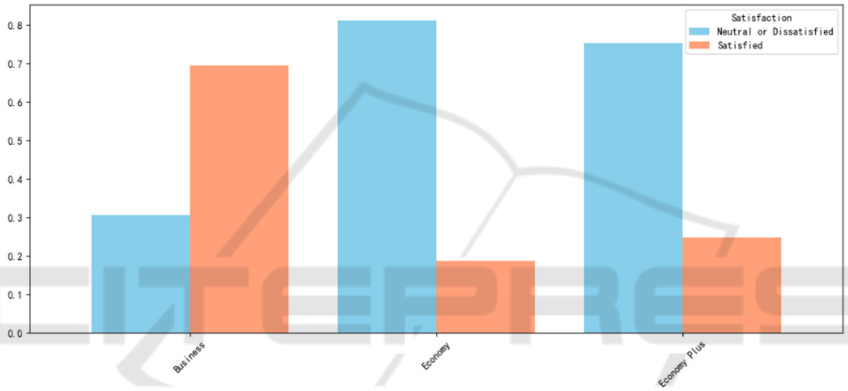

Figure 17: Distribution chart of overall passenger satisfaction for different flight destinations.

their precious time, such as flight delays and complicated check-in procedures.

Therefore, it is recommended that airlines continue to provide high-end services to business travelers while striving to ensure that flights take off and land on time to reduce the inconvenience and pressure on business travelers' schedules. And improve business travelers' evaluation of the convenience of flight times through effective operations and on-time flight execution; At the same time, attention should also be paid to investing in personal passenger services and cultivating passenger loyalty to improve the situation where the number of neutral or dissatisfied passengers far exceeds that of satisfied passengers.

4.5 Analysis of Differences in Satisfaction among Passengers of Different Cabin Classes

As shown in the Figure 17 and Figure 18.

As can be seen from the above Figure 17 and 18, the overall satisfaction of business class passengers is much higher than that of economy class and premium economy class passengers, while the overall satisfaction of premium economy class passengers is slightly higher than that of ordinary economy class passengers. In contrast, business class passengers also have higher demands for convenience in departure and arrival times than economy class passengers.

Therefore, the average rating of the corresponding items is lower. The average rating of the boarding gate location and on-board Wi-Fi service is similar to that of economy class passengers. Therefore, it is recommended to improve the quality of in-flight wireless networks, optimize boarding gate guidance services or provide shuttle buses, and strengthen services for economy class passengers to solve the problem of high proportion of neutral or dissatisfied passengers.

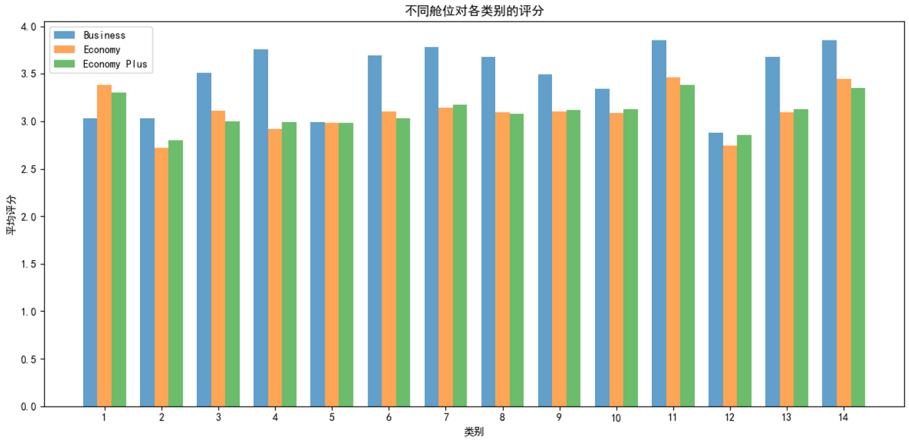

Figure 18: Histogram of average passenger ratings for different flight destinations.

5 CONCLUSIONS

The article selects the XGBoost model, combine with SHAP model prediction and analysis to find that online boarding, legroom services and in-flight Wi-Fi services are the most important to be focused on improvement. At the same time, this article also puts forward relevant improvement suggestions for these services. Such as optimizing advance seat selection and online check-in procedures, providing wider seat space and more comfortable seating experience, regular maintenance and upgrade of in-flight wireless network facilities, etc.

This article analyzes the differences in satisfaction among different types of passengers in five aspects. In terms of age, satisfaction among younger and older age groups is lower, mainly reflected in online check-in, in-flight services and in-flight entertainment. Therefore, it is recommended that airlines understand the preferences of passengers of all ages and provide personalized services. In terms of gender, there is no significant difference in the overall satisfaction of male and female passengers, and the ratings of various services are similar. Therefore, it is determined that gender has no significant impact on passenger satisfaction. In terms of customer types, the satisfaction of returning passengers is significantly higher than that of first-time passengers, specifically in terms of legroom service and online check-in. Therefore, it is recommended that airlines improve more detailed boarding guides and instructions on facilities use to help first-time passengers better adapt to the flight. Regarding the purpose of the flight, the satisfaction of business travelers is much higher than that of individual travelers, and business travelers also rate most services higher than individual travelers. It is recommended that airlines strive to ensure the punctuality and convenience of flights and enhance their investment in services for individual travelers. In terms of cabin class, the satisfaction level of business class passengers is obviously the highest, followed by the satisfaction level of premium economy class passengers, and the lowest satisfaction level of economy class passengers. This is also consistent with the actual situation. It is recommended to focus on optimizing the in-flight Wifi service with lower ratings.

Overall, passengers' overall satisfaction is low, with the convenience of online reservations, gate location and in-flight Wifi receiving the lowest ratings among various services. It is recommended that airlines optimize services based on actual conditions, especially in categories with low ratings and high importance, to improve passenger satisfaction.

REFERENCES

Alsayat, A, 2023. Customer decision-making analysis based on big social data using machine learning: a case study of hotels in Mecca. *Neural Comput Applic*. 35, 4701–4722.

Chung, S., 2022. The influence of CSR on airline loyalty through the mediations of passenger satisfaction, airline brand, and airline trust: Korean market focused. *Sustainability*, 14(8), 4548.

Guan, G., Liu, D., Zhai, J. 2022. Factors influencing consumer satisfaction of fresh produce E-commerce in the background of COVID-19—A hybrid approach based on LDA-SEM-XGBoost. *Sustainability*, 14(24), 16392.

Garmendia Mora, J. C., Tavera Morales, D. A, 2021.Aerolíneas bajo costo y convencionales: la satisfacción y lealtad de sus clientes. *Revista Universidad & Empresa,* 23(41), 1-26.Mateja K, 2017. An investigation of customer satisfaction with low-cost and full-service airline companies, *Journal of Business Research*,80, 188-196

Hyunwoo P, 2022. Hyunju J, The Dynamics of Customer Satisfaction Dimension based on BERT, SHAP, and Kano Model, IFAC-PapersOnLine, 55, 10, 2384-2389.

Sao, A., Sharma, N., Singh, S., Vishwas Yelikar, B., Bhardwaj, A. 2023. Examining Challenges In The Adoption of Big Data In Health Care Institutions And Its Impact On Patients Satisfaction: An empirical study in Delhi, India: - Global Conference on Emerging Technologies, Business, Sustainable Innovative Business Practices, And Social Well-Being. *Asia Pacific Journal of Health Management*, 18.

Sudhakar, S., Gunasekar, S, 2020. Examining online ratings and customer satisfaction in airlines. *Anatolia,* 31, 2, 260–273.

Thekkoote, R. 2022. Understanding big data-driven supply chain and performance measures for customer satisfaction. *Benchmarking*, 29(8), 2359-2377.

Xu, Y., Wu, G., Chen, Y, 2022. Predicting patients' satisfaction with doctors in online medical communities: An approach based on XGBoost algorithm. *Journal of Organizational and End User Computing,* 34(4), 1-17.