Journal of Air Transport Management 125 (2025) 102795

Contents lists available at ScienceDirect

Journal of Air Transport Management

journal homepage: www.elsevier.com/locate/jairtraman

A variant bagging forecasting framework for customer churn in airline

Qiang Li [a], Yuangang Li [b], Ranzhe Jing [c,*]

[a] School of Management, Jiangsu University, Jiangsu, Zhenjiang, 212013, PR China
[b] Faculty of Business Information, Shanghai Business School, Shanghai, 200235, PR China
[c] School of Information Management & Engineering, Shanghai University of Finance and Economics, Shanghai, 200433, PR China

ARTICLE INFO

Keywords:
Customer churn
Quality of service
Hedonistic and utilitarian factors
Bagging
Data mining
Airline industry

ABSTRACT

The goal of this study is to forecast customer churn and analyze the influence of quality of service on customer churn in airline industry. Following a multifactor approach, a variant Bagging forecasting framework is proposed to mine the inner patterns of customer churn. A probabilistic sampling approach is embedded in the developed model simulating the customer churn probabilities. The airline customer feedback data by airline carriers in the U.S. was used to train the prediction model. The results indicate the accuracy for predicting customer churn is 96 %, and the most important factors for customer churn are in-flight entertainment, seat comfort, and type of travel. We also investigated the effects of service quality (both hedonistic and utilitarian factors) on various customer groups and discovered that improving hedonistic service quality can effectively reduce customer churn.

1. Introduction

The aviation industry has become more competitive while also developing airline firms as a result of the quick increase in demand for air travel (Belobaba et al., 2015; Calisir et al., 2016). In such a cutthroat environment, maintaining a high level of passenger satisfaction is regarded as a crucial business advantage (Hussain et al., 2015; Noviantoro and Huang, 2022). Numerous studies have been done to demonstrate the link between customer happiness and customer loyalty and results indicate high levels of customer satisfaction encourage travelers to book future flights with the same airline or to refer it to others (An and Noh, 2009; Forgas et al., 2010; Chen and Hu, 2013). On the other hand, low customer satisfaction can harm a business's reputation and image in addition to contributing to issues with customer churn (Blodgett and Li, 2007). According to Athanassopoulos' (2000) findings, acquiring a new customer costs a business between $300 and $600, which is almost five times as much as keeping an existing clientele (Athanassopoulos, 2000). Therefore, each company's primary goal is to lower customer churn by attempting to raise customer satisfaction (Agarwal and Gowda, 2021). The most efficient way to stop customer turnover is to establish precise churn prediction models, identify possible lost clients beforehand, and create precise retention strategies (Sharma and Panigrahi, 2011; Gordini and Veglio, 2017; De Caigny et al., 2020).

Many academics domestically and internationally have started to use these technologies in the field of predicting customer churn in recent years due to the rapid development of machine learning and deep learning technologies, particularly their considerable benefits in prediction (Huang et al., 2012; Vafeiadis et al., 2015). Tsai and Lu (2009) used artificial neural network (ANN) and self-organizing mapping (SOM) for customer churn prediction while conducting research abroad. Where the training set's unrepresentative data is filtered out using ANN. The output results are subsequently fed back into the SOM to build the prediction model. The study's findings demonstrate that an ANN + SOM combo performs more accurately than a single neural network. In order to increase the model prediction accuracy, He et al. (2014) integrated a random sampling strategy with a support vector machine model to account for the imbalance between churned and non-churned customer samples in the sample data.

To theoretically guide businesses in creating effective customer retention strategies, firms must simultaneously establish a high-precision model for predicting customer churn and identify and study the major elements that influence customer churn. The current research is more focused on increasing the model's predictive accuracy; no more in-depth research has been done on the identification of critical variables or the impact of service quality on customer turnover. As a result, there are three main areas of emphasis in this work. First, based on machine learning algorithms and actual flight customer data, we build a customer churn prediction model to achieve early warning of potential customer churn. Secondly, based on the constructed prediction model,

* Corresponding author.
E-mail address: jing.ranzhe@sufe.edu.cn (R. Jing).

https://doi.org/10.1016/j.jairtraman.2025.102795
Received 29 May 2024; Received in revised form 14 February 2025; Accepted 29 March 2025
Available online 4 April 2025

Table 1
Summary of factors and their statistics.

Factors	Description	value
Satisfaction	Customer satisfaction with airlines	satisfied(54.7 %) unsatisfied(45.3 %)
Age	Actual age of passengers	≤20 (10.8 %) 20 - 40 (41.0 %) 40 -60 (40.5 %) ≥60 (7.7 %)
Gender	Passenger's gender	Male(50.74 %) Female(49.26 %)
Type of travel	Personal travel or business travel	Personal travel (30.9 %) Business travel (69.1 %)
Flight class	Business, economy or economy plus	Business (47.9 %) Economy (44.9 %) Economy plus (7.2 %)
Customer type	Member or not	Member(84.6 %) Non-member (15.4 %)
Flight distance	The flight distance of this journey	250-5000 mile
Flight delays	Minutes delayed	0–3176 min
Seat comfort	Satisfaction level of Seat comfort	Rating from 1 to 5
Departure/Arrival time convenient	Satisfaction level of Departure/ Arrival time convenient	Rating from 1 to 5
Food and drink	Satisfaction level of Food and drink	Rating from 1 to 5
Gate location	Satisfaction level of Gate location	Rating from 1 to 5
Inflight wifi service	Satisfaction level of the inflight wifi service	Rating from 1 to 5
Inflight entertainment	Satisfaction level of the Inflight entertainment	Rating from 1 to 5
Online support	Satisfaction level of Online support	Rating from 1 to 5
Ease of Online booking	Satisfaction level of Ease of Online booking	Rating from 1 to 5
On-board service	Satisfaction level of On-board service	Rating from 1 to 5
Leg room service	Satisfaction level of Leg room service	Rating from 1 to 5
Baggage handling	Satisfaction level of Baggage handling	Rating from 1 to 5
Checkin service	Satisfaction level of Checkin service	Rating from 1 to 5
Cleanliness	Satisfaction level of Cleanliness	Rating from 1 to 5
Online boarding	Satisfaction level of Online boarding	Rating from 1 to 5

the importance level of different factors is analyzed to explore the key customer churn influencing factors. Finally, 1000 airline customers are randomly screened to study the impact of service quality changes on the number of lost customers, and to make customer churn prevention decisions for the company.

2. Problem definition and model

In this section, we introduce the dataset used for this study, define the key terms related to customer churn, and describe the research model. The dataset includes various factors that are hypothesized to affect customer churn. We first describe the dataset and the related service attributes in Table 1, which serves as the foundation for the customer churn model. Each factor from the dataset plays a significant role in understanding the drivers of customer satisfaction and ultimately, customer turnover.

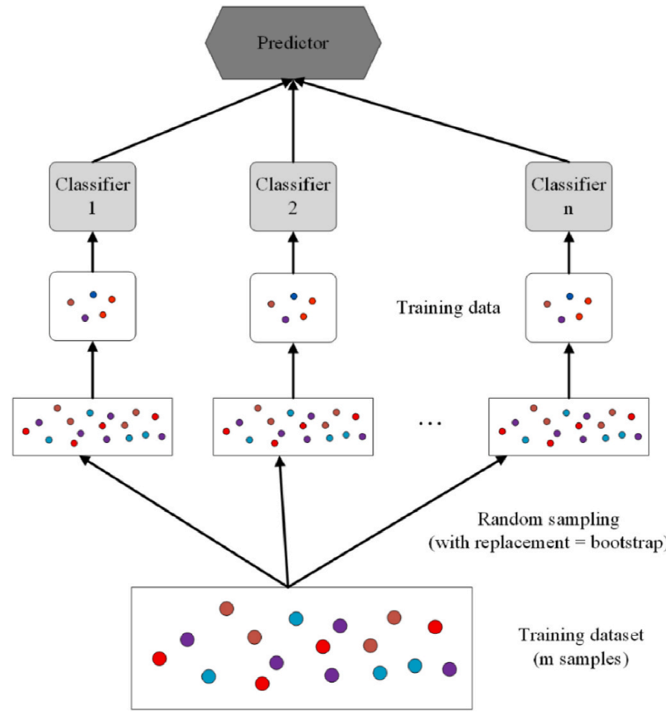

Fig. 1. Framework of Bagging algorithm.

2.1. Datasets and definitions

This part builds a customer churn model and examines the effect of service quality on customer churn using a sample of 129,880 feedback messages from the US airline after its passengers have finished their travels (Klein, 2015). The study's dataset includes data on the respondents' ages, genders, flight classes, customer types, and satisfaction levels in addition to a total of 14 service attributes in the areas of online service, airport service, and in-flight service. Service attributes were evaluated on a five-point scale, with scores ranging from 1 to 5 corresponding to "very bad," "somewhat bad," "neutral," "somewhat good," and "very good." Table 1 provides a detailed description of the datasets.

Customer churn, as described by Kamakura et al. (2005), refers to the act of a customer leaving or switching to another business. According to Shahin and Zairi (2009), approximately 96 % of dissatisfied customers do not complain about airline services, and 90 % of these silent complainers never return, leading to a high attrition rate. Building on this research, we hypothesize that non-member and unsatisfied airline customers in our sample of 129,880 entries have an 86.4 % probability of churning (calculated as 96 % × 90 % = 86.4 %).

Positive customers: Members who are satisfied with the airline's services.

Remedial customers: Members who are dissatisfied with the airline's services.

Potential customers: Non-member consumers who are pleased with the airline's services.

Lost customers: Non-member clients who are displeased with the airline's services.

According to Shahin et al., (2009), non-member customers who are dissatisfied with airline services do not necessarily churn but have an 86.4 % chance of churning. As a result, to determine the sample of churned customers in the subsequent modeling analysis, we sampled the churn-level customer group with probability $p = 0.864$.

2.2. Forecasting customer churn using the bagging variant model

We must sample the churn-level customer population with $p = 0.864$

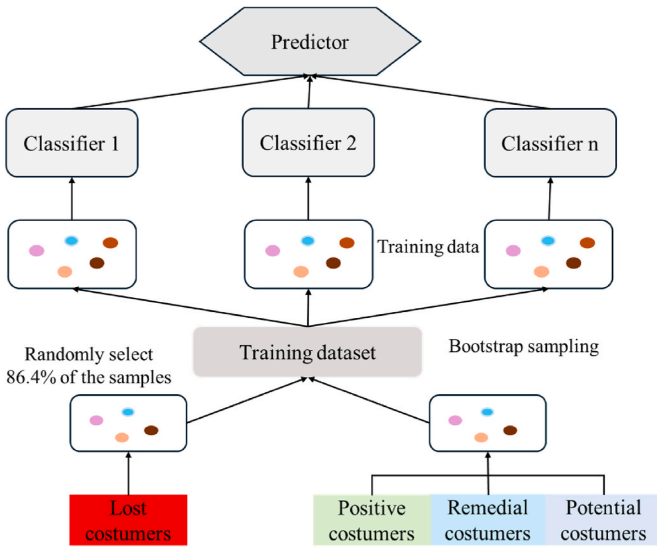

Fig. 2. Variant of Bagging algorithm.

as the probability while performing training sample selection since the percentage of customers who really churn in the churn-level customer population is approximately 86.4 %. Due to this, the Bagging technique (Breiman, 1996) used to build the prediction model is the main emphasis of this section.

The most well-known example of a parallel integrated learning approach is the bagging algorithm, and Fig. 1 depicts its fundamental operation. Based on the process in Fig. 1, the Bagging method is used to construct and combine n weak learners in order to fulfill the learning goal. The booststrap sampling approach is used to generate training samples for the n weak learners. The self-sampling method's basic operation entails selecting one sample at random from the original training set of m samples, placing it back into the sampling set m times in a sequence, and finally obtaining the training set of m samples.

In this study, we adjust the sample approach depicted in Fig. 1 and suggest a modified version of the Bagging algorithm as depicted in Fig. 2, because the proportion of customers who actually churn in the churn-level customer group is approximately 86.4 %.

In contrast to the conventional Bagging technique, we split the m training samples into two sub-datasets based on the categories of customers: those that churn and those who are positive, improving, and remaining clients. We use a random sampling method to select 86.4 % of the data from the churn-level customers to represent the churned customer sample because there is an 86.4 % probability that the customer group will actually experience churn, while the self-service sampling method is still used for sampling in the other dataset. The data acquired by combining the two sampling techniques serve as the weak learner's training sample.

We use the Logistic Regression (DeMaris, 1995), K-nearest neighbor (Kramer, 2013), Support vector machine (Noble, 2006), and Decision tree models (Brijain, 2014), respectively, for the selection of weak learners. The traditional random forest model is the bagging algorithm, which is built with a decision tree as the weak learner. In order to output the final prediction results, we choose the voting method for combining strategies, meaning that a category with more than half of the votes is predicted to be that category.

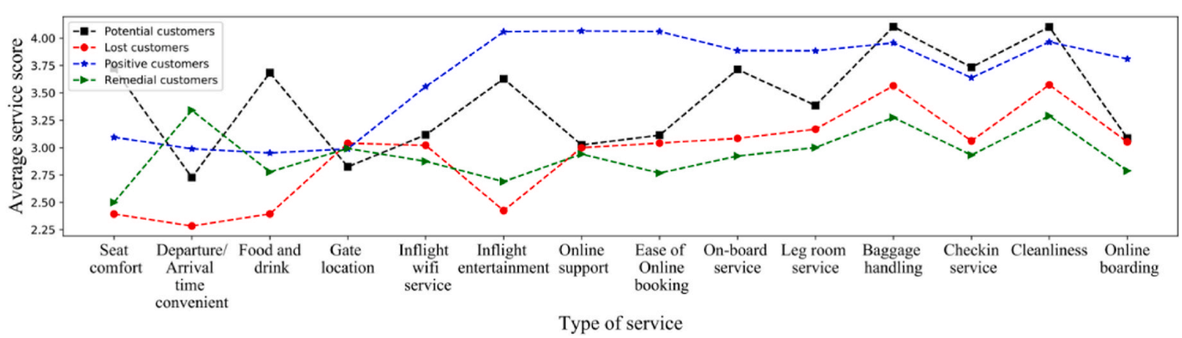

Fig. 3. The relationship between customer type and factors.

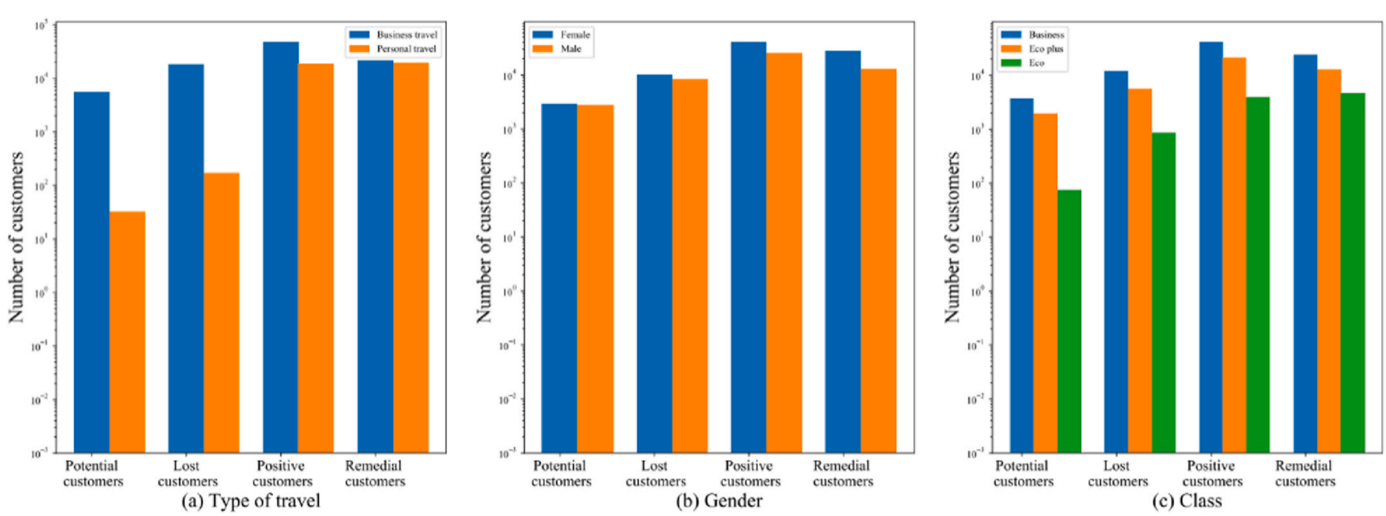

Fig. 4. Four major categories of customer distribution under different customer groups.

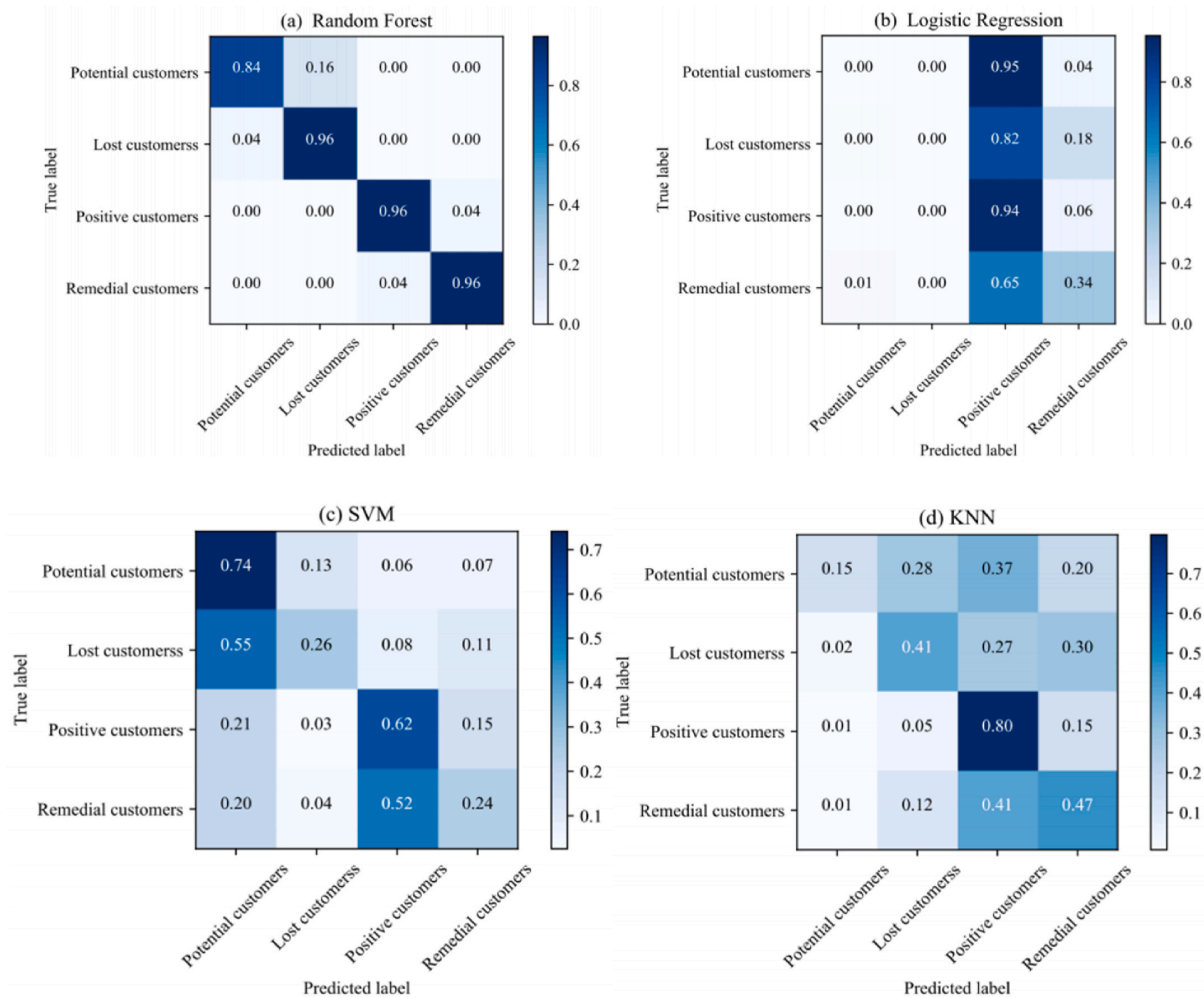

Fig. 5. The prediction results for different models.

3. Data analysis and prediction results

The basic data properties are described in this part, along with the model's predictions using various classifiers. Fig. 3 displays the ratings for 14 service criteria for four different costumer types: Positive customers, Remedial customers, Potential customers, and Lost customers.

According to the study's findings, Positive customers and Potential consumers have the highest service ratings. This is largely because both groups of customers are pleased with the services offered by the airline and as a result, have higher opinions of those services. Customers who received remedial service from an airline give the company's service a lower rating than Lost customers. The primary cause of this is that member consumers have higher service expectations than non-member customers, which leads to a worse service rating when the airline falls short of their fundamental standards.

Fig. 4 breaks down the statistics on customer attrition according to the types of passenger travel (personal vs. business), gender (male vs. female), and seat class (business class, economy class, or discount economy class). Fig. 4's findings demonstrate that, on the contrary, the probability of churn is lower for the personal trip client group than the business travel group. There is no discernible difference between them and customer churn in terms of gender features. And when it comes to seat class, the business class client segment has the highest attrition rate while the discount economy segment has the lowest churn rate.

According to the findings in Fig. 4, the business travel and business class customer groups experience a higher rate of churn than the individual travel and economy class customer groups. Therefore, in order to reduce customer churn, airlines must pay more attention to this group's input.

In the next investigation, we used logistic regression (LR), K-nearest neighbor (KNN), support vector machine (SVM) models as weak classifiers and random forest (RF) models for customer churn prediction analysis, respectively. The confusion matrix of the prediction results of the four models is shown in Fig. 5.

The results of the study in Fig. 5 show that the best prediction performance is achieved using Random Forest algorithm. Its prediction accuracy for the four types of customers Positive customers, Remedial customers, Potential customers and Lost customers reached 84 %, 96 %, 96 % and 96 %, respectively. The logistic regression as a weak classifier was the least effective. As a result, the random forest model for customer churn prediction is quite accurate at predicting probable lost clients. Taking the experimental results in Fig. 5(a) as an example, when random forest is used for churn prediction, about 96 % of churned customers are successfully predicted, while the model only predicts about 4 % of

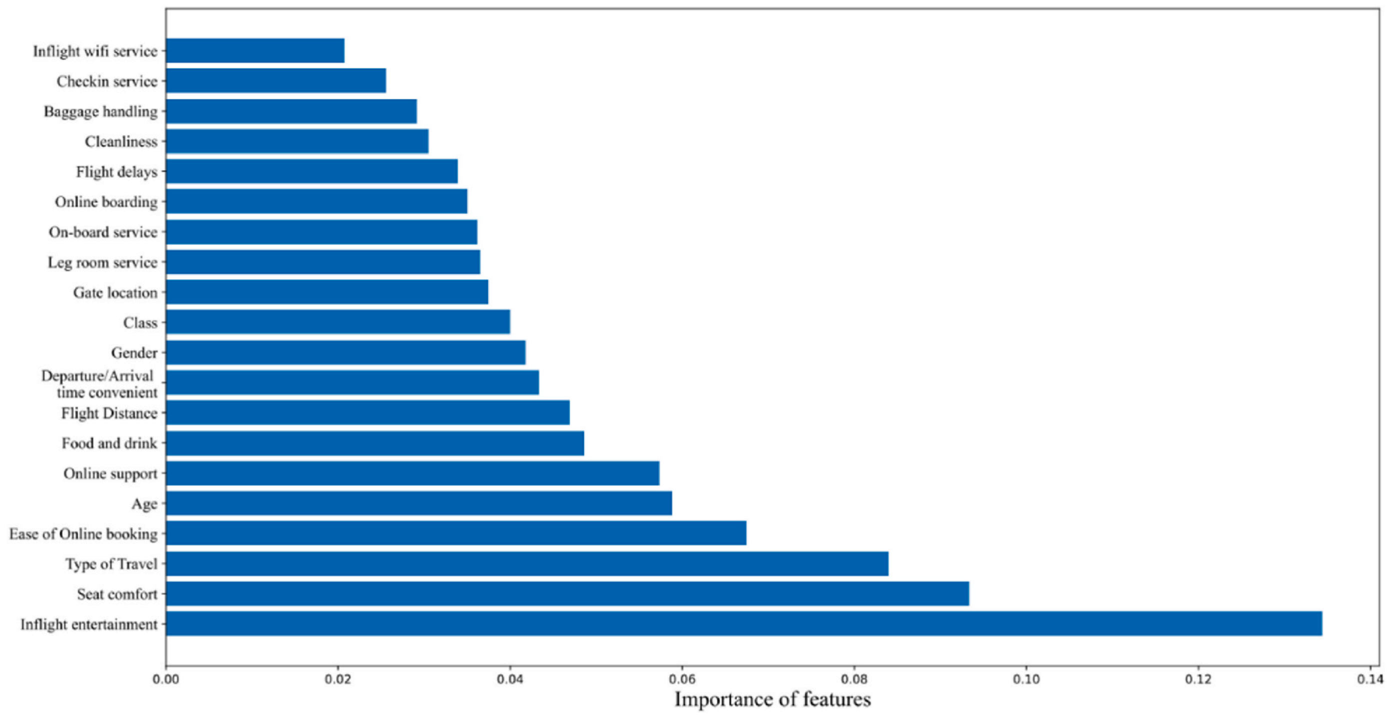

Fig. 6. Ranking the importance of features affecting customer churn.

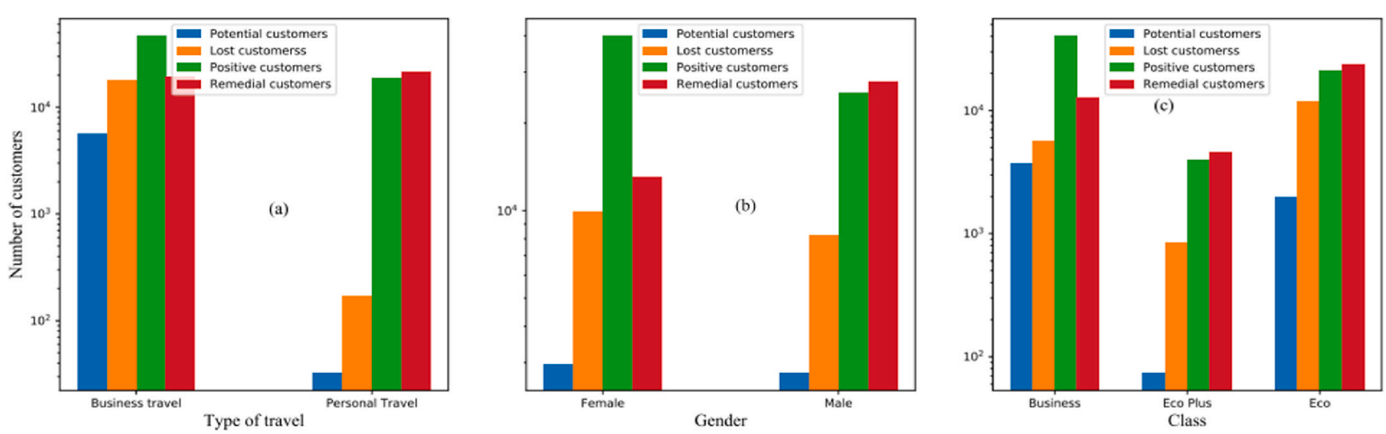

Fig. 7. Relationship between type of travel, gender, class and customer type.

churned customers as retention level customers. Based on the results in Fig. 5, we can find that the model we constructed has high prediction accuracy and is a good reference for correctly identifying potential churned customers, further validating the feasibility and accuracy of constructing a customer churn prediction model.

4. Identification of important features affecting customer churn

Another advantage of the random forest model is that it can automatically output the importance scores of each feature. For the customer churn problem, the importance of each influencing factor is shown in Fig. 6. Where the horizontal coordinates indicate the importance scores of each variable and the vertical coordinates indicate the input variables of the model. The findings in Fig. 6 show that three variables such as Inflight entertainment, Seat comfort and Type of Travel have the highest impact on customer churn with importance scores of 0.13, 0.09 and 0.08 respectively, while Inflight Wifi service and Gate Location have the least impact on customer churn with 0.021 and 0.024 respectively. Inflight Wifi service and Check in service have the least impact on customer churn with 0.02 and 0.03 respectively. Other category features, such as gender and class, have important scores of 0.042 and 0.040, respectively, ranking 10 and 11 among all features.

To investigate the impact of potential factors on customer churn, we investigate the impact of the top five important factors on customer churn, as well as the impact of category features such as type of travel, gender, and class. Fig. 7 illustrates the relationship between customer type and type of travel, gender, and class while the influence of the top five factors is shown in Fig. 8.

From the results of Fig. 7, we found that the number of positive level customers and potential customers in the business travel passenger group is much higher than the personal travel passenger group, but on the other hand, the total number of churn in the business travel group is also 17,911, which is about 91 times more than the 196 in the personal travel group. Therefore, airlines should pay more attention to the business travel customers to reduce their churn rate. In terms of seat class, both business class and economy class customers have higher

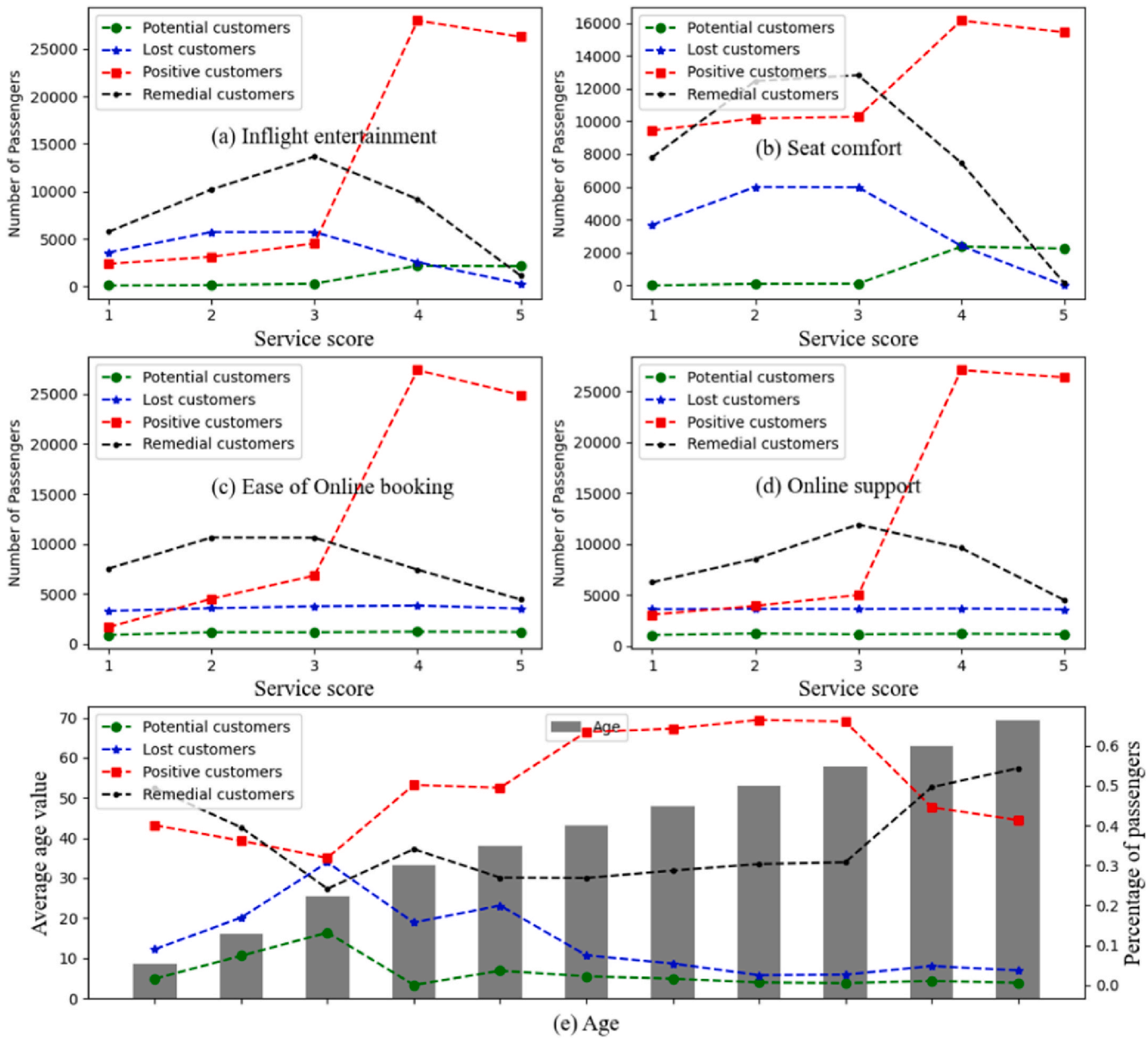

Fig. 8. The relationship between the top five factors and customer type.

churn rates, but business class customers have more market potential overall, mainly reflect in the number of positive and potential customers. From a gender perspective, there has no significant difference between male and female customer segments in general.

The impact of the top five factors (Inflight entertainment, Seat comfort, Type of Travel, Ease of Online booking and Age) on customer churn are illustrated in Fig. 8, where the X axis represents the rating of service quality, and the Y axis is the number of customers.

From the results in Fig. 8, we discovered a non-linear relationship between the top five factors and customer churn. For instance, when the in-flight entertainment rating score reaches three points, the amount of customer churn reaches its greatest level, totaling roughly 5723 customers. Another interesting result presented in Fig. 8 is that the service score mainly affects the positive customers, and the customer churn will not always decrease with the increase of service score. As shown in Fig. 8 (a and b), when the service score rises from three to five, the positive customers increase from 1134 to 26270 and 9446 to 15433, respectively, and the percentage of positive consumers reaches the highest values when the rating approaches four points. For customer churn issues, Fig. 8 indicates that customer churn is not negatively correlated with service quality (Gao et al., 2021). When the service quality for in-flight entertainment and seat comfort exceeds 3, customer attrition decreases. In terms of ease of online booking and online support, service quality and customer churn are not strongly correlated.

Fig. 8 (e) explores the impact of age on customer churn and results indicate that younger customer groups-primarily those aged 20 to 30, have considerably greater rates of customer churn, however there is no statistically significant correlation between customer churn and age among those who are more than 40 years old. The results presented in Fig. 8 also show a service score of three is a critical turning point for customer churn and the positive customer and remedial customer are more likely to be affected by service quality compared with the potential customer and lost costumer.

To further investigate the correlation between the top five factors and customer churn, we apply the trained random forest model to simulate the impact of these factors on customer churn. The

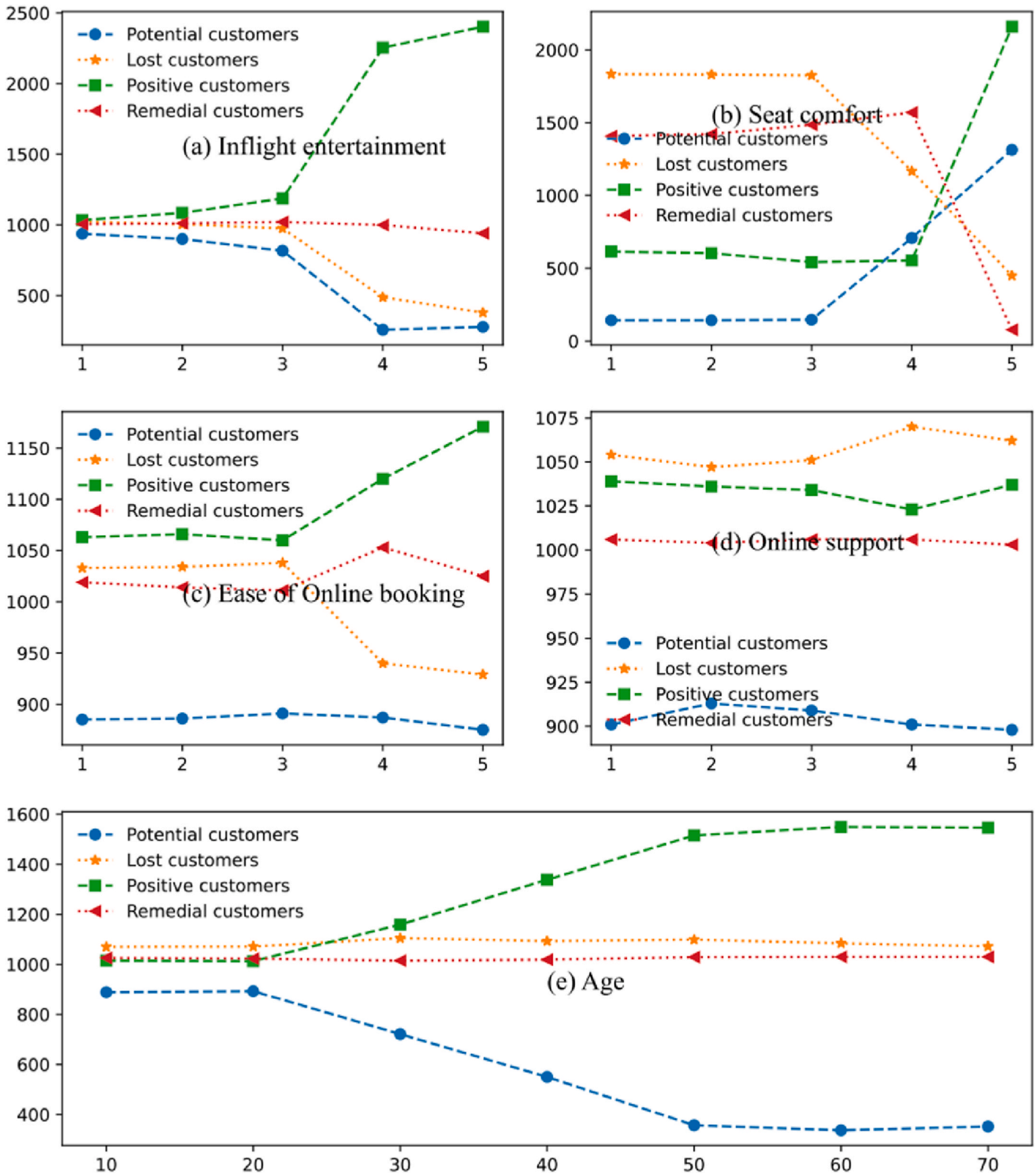

Fig. 9. The changes of the number of customers in the four customer groups with the increase of inflight entertainment, seat comfort, ease of online booking, online support, and age.

fundamental modeling steps to enhance how the top five factors affect customer churn are displayed below.

i. Picking 1000 consumers each at random from positive level, potential level, retention level and churn level samples.

ii. Maintaining the other qualities constant for the 4000 sample data points and increasing the score of the five factors from 1 to 5 respectively.

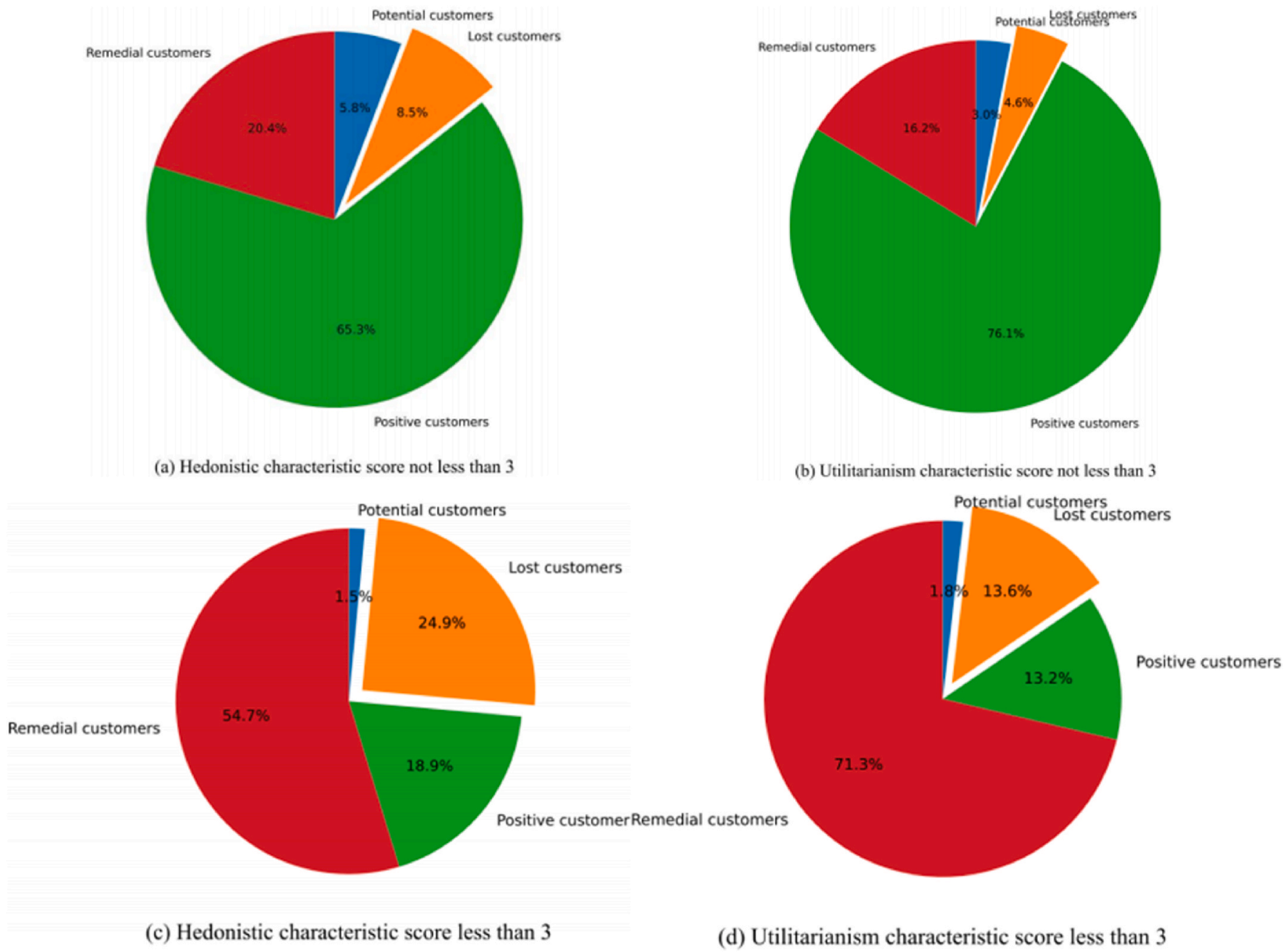

Fig. 10. The service score of hedonistic and utilitarian factors *vs* the type of customer.

iii. Exhibiting the change curves of the number of customers in the four customer groups using the trained random forest model as the prediction model for the 4000 samples.

Based on the results in Fig. 9, we found online support and age have a minimal impact on customer churn, whereas Inflight entertainment, Seat comfort, and Ease of online booking significantly affect customer churn. Additionally, we observed that a significant change in customer churn occurs only when service quality increases from 3 to 5; an increase from 1 to 3 does not alter customer churn. This suggests that customers have an expected level of service, and only when the service quality exceeds these expectations does an improvement influence customer churn.

5. The influence of hedonic and utilitarian factors on customer churn

To investigate the impact of hedonistic and utilitarian factors on customer churn, we categorize all factors into two types: hedonistic and utilitarian factors and reveal their relationship with customer churn in this paper.

Hedonism refers to services that provide customers with comfort and entertainment, and in this paper, our hedonism services include features such as Seat comfort, Food and drink, Inflight Wifi service, Inflight entertainment, Leg room service, Cleanliness. Utilitarian services are those that make it easier for customers to travel, such as Departure/Arrival time convenient, Gate location, Online support, Ease of Online booking, On-board service, Online boarding, Baggage handling, Check-in service. In this paper, we use a service score of three as the cut-off point, i.e., neutral, to examine the distribution of different customer types when hedonistic and utilitarian factors are greater than or less than three, Fig. 10 depicts the results.

From the results in Fig. 10, we discovered when the service score of hedonistic factors is not less than three, only 8.5 % of customers will lose. However, if the hedonistic score was less than three, the percentage of lost customers increased from 8.5 % to 24.9 %, nearly tripling, indicating the importance of hedonistic factors on customer churn. Compared with the hedonistic factors, when the utilitarianism score changed from not less than three to less than three, customer churn has also increased by about three times. However, when contrasting hedonism to functionalism, we notice the hedonistic trait caused a churn rate of 24.9 % when both the hedonistic and utilitarianism features were rated lower than 3, which was nearly twice as high as the churn rate in the event of low utilitarianism ratings. Besides, the hedonistic kind also causes increased churn when both are rated higher. As a result, we continue to investigate how hedonistic and utilitarian factors affect customer churn, particularly how they affect various customer types. We simulate the change of hedonistic and utilitarian factors on customers using the proposed prediction model, which is similar to the procedure in Section 4.

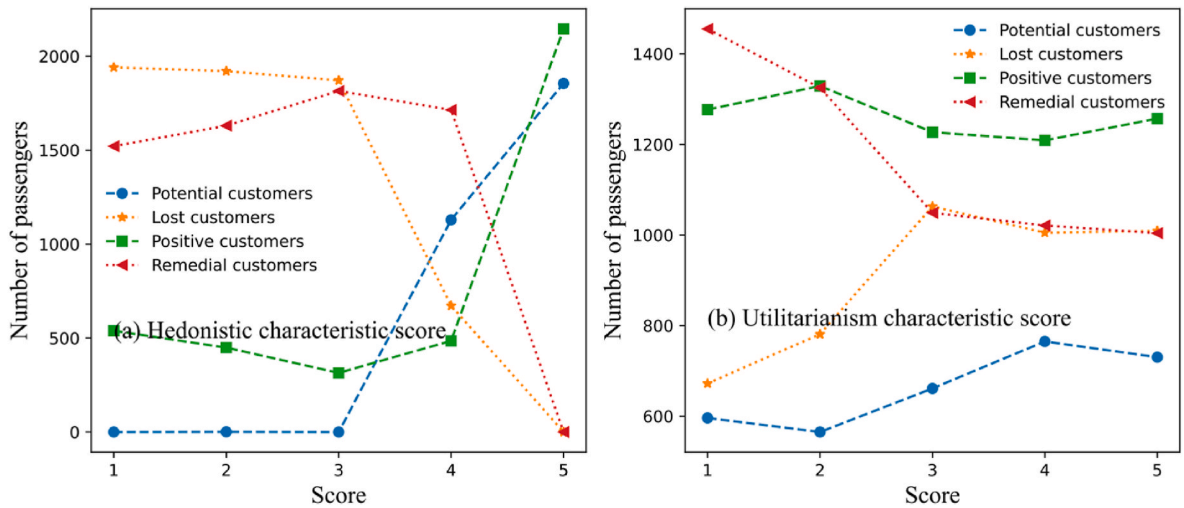

Fig. 11. The change of hedonistic and utilitarian factors on customers.

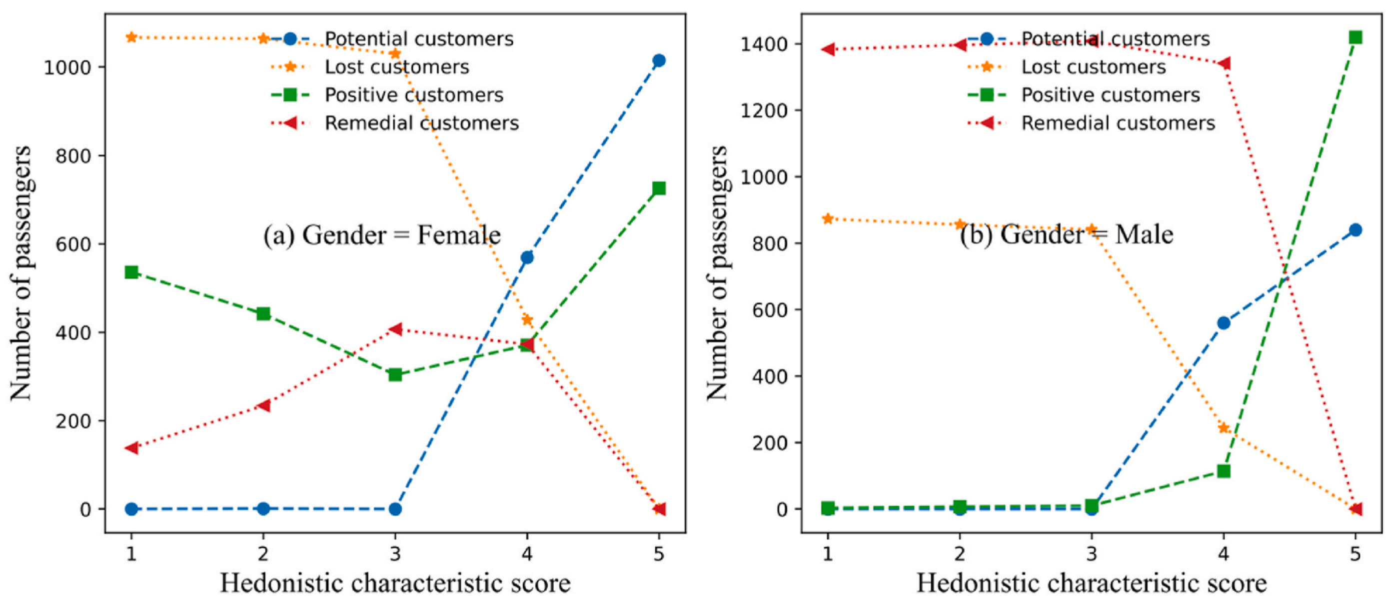

Fig. 12. Hedonistic factors' effects on customer churn among various gender groups.

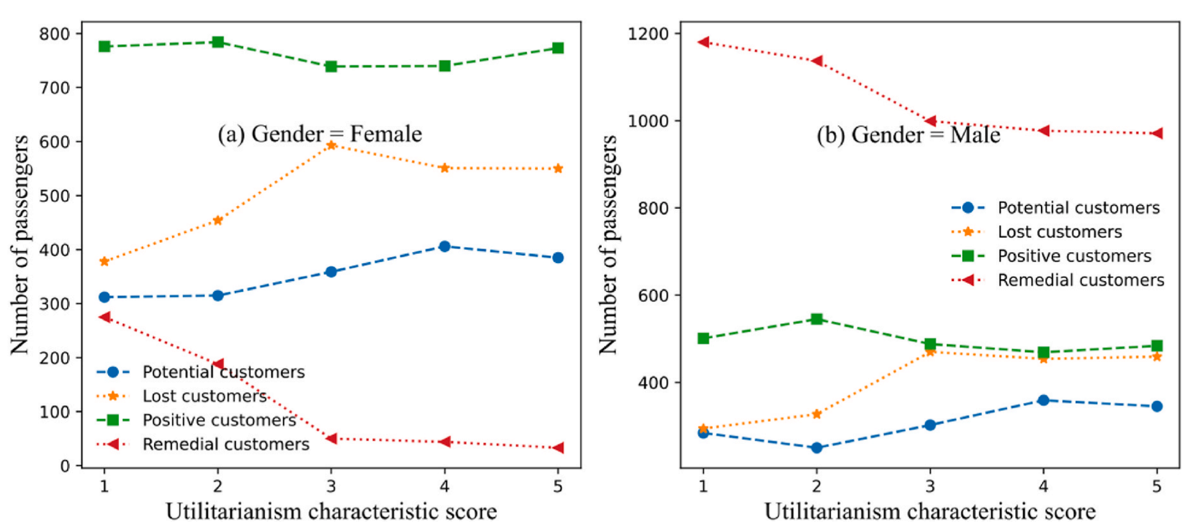

Fig. 13. Utilitarianism factors' effects on customer churn among various gender groups.

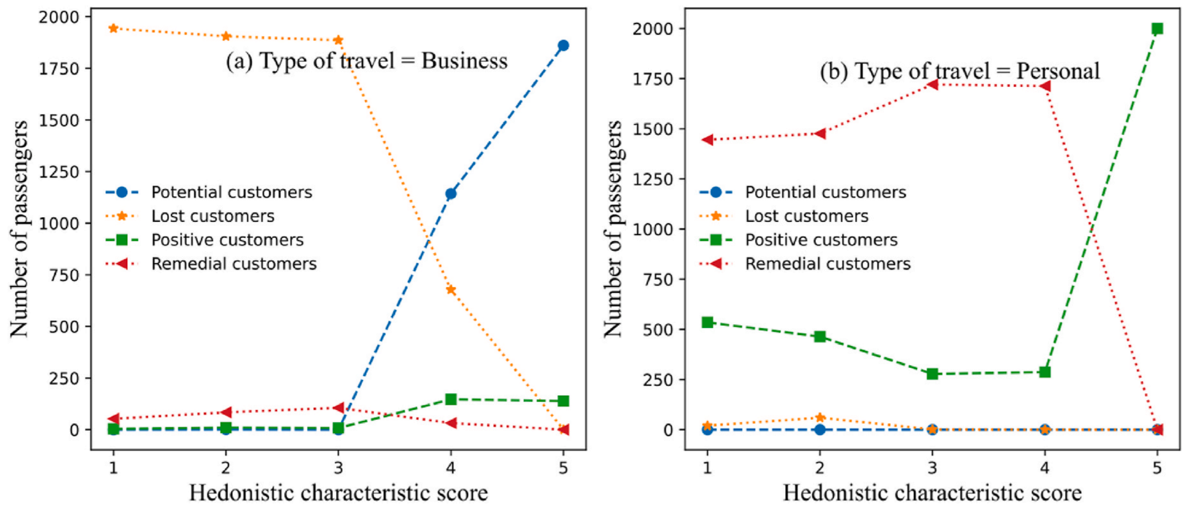

Fig. 14. Hedonistic factors' effects on customer churn among various type of travel groups.

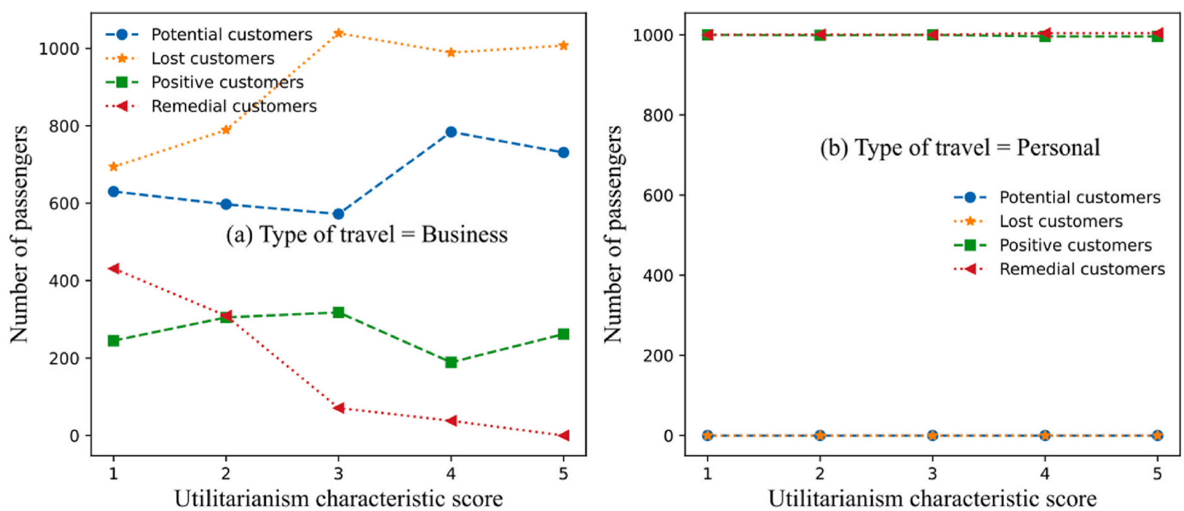

Fig. 15. Utilitarianism factors' effects on customer churn among various type of travel groups.

i Picking 1000 consumers each at random from positive level, potential level, retention level and churn level samples.
ii Maintaining the hedonistic/utilitarian factors constant for the 4000 sample data points and increasing the score of the utilitarian/hedonistic factors from 1 to 5 respectively.
iii Exhibiting the change curves of the number of customers in the four customer groups using the trained random forest model as the prediction model for the 4000 samples.

The influence of hedonistic and utilitarian factors on customers is shown on Fig. 11.

From the results in Fig. 11(a), we found that hedonic factors only begin to have an impact on customer attrition if the service score is greater than three. The majority of customers are essentially in a lost and remedial condition when the hedonic variable score is below three. The number of lost and remedial clients dramatically decreases and changes to positive and potential customers as the hedonic variable score rises over three. Fig. 11(b) demonstrates that the utilitarian factors have a larger impact on the remedial and lost customer groups than on the positive and potential customer groups. In contrast hand, we see an unexpected negative tendency in customer churn when evaluations of functional factors continue to rise. We discovered that the component convenient departure/arrival times is the primary factor contributing to this finding by further studying the association between functional characteristics and customer attrition. Longer waiting times between flights are indicated by a more convenient transfer time. Additionally, the prolonged waiting time lowers consumer satisfaction, which worsens turnover issues. The findings in Fig. 11 illustrate that hedonic factors play a higher role in customer churn compared to utilitarian factors.

We examined the impact of changes in hedonic and utilitarian characteristics on the number of churned customers to further investigate the impact of hedonic and utilitarian characteristics on different categories of customers (e.g., gender, type of travel). The study's findings are depicted in Figs. 12–16.

Fig. 12 illustrates the impact of hedonistic factors' effects on customer churn for different gender groups. The results indicate when the hedonistic score over three, churn drops significantly for both male and female customer segments. When the hedonistic score is 5, there is essentially no customer churn.

Fig. 13 illustrates the impact of utilitarianism factors' effects on customer churn for different gender groups. The results indicate utilitarianism factors have a positive trend and have less of an effect on customer attrition compared with the hedonistic factors. To lower customer churn, airlines should focus more on hedonistic service quality than utilitarianism service quality.

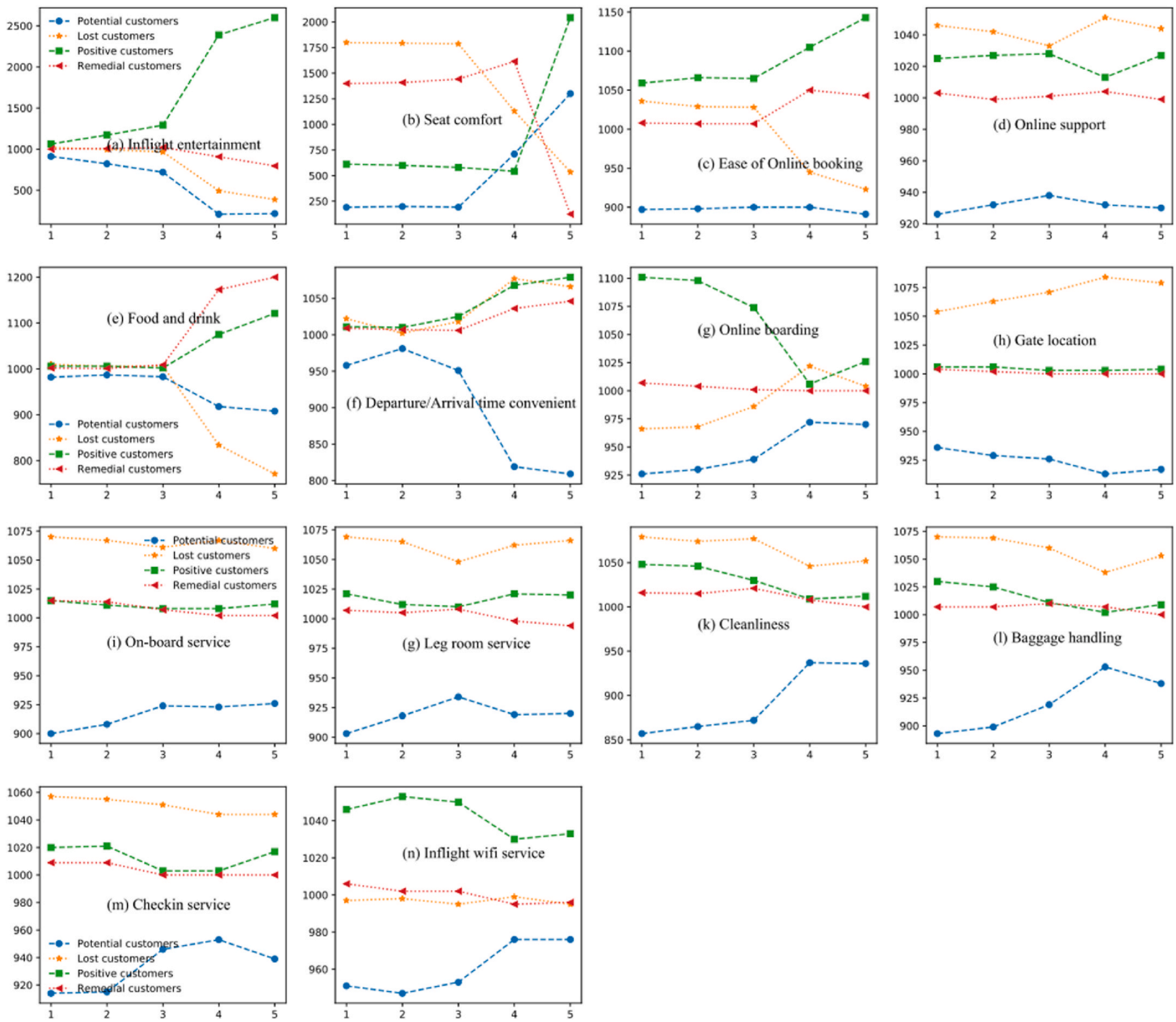

Fig. 16. The influence of all hedonistic and utilitarianism factors on customer churn.

Fig. 14 describes the effects of hedonistic factors' on customer churn for business and personal travelers. From the results on Fig. 14, we found the hedonistic factors mainly affect the business traveler, especially the service score over three. For the individual travel customer group, the hedonistic variables have a particularly small impact. As a result, improving the level of hedonic service quality in business class can effectively reduce the problem of customer churn for airlines.

Fig. 15 descripts the effects of utilitarianism factors' on customer churn for business and personal travelers, and the results also surface utilitarianism variables only affect business travelers.

Hedonistic and utilitarianism factors' effects on customer churn among various type of travel groups indicate airline companies should focus on improving the quality of their services for the enjoyment of business travelers rather than on functional services. The first is that improving the quality of hedonistic services can effectively reduce customer churn, whereas functional services have little or no impact on customer churn, as shown in Fig. 16. Personal travelers, on the other hand, are less sensitive to hedonistic and functional services, and changes in service quality have little impact on their churn.

6. Conclusion

Customer churn is a topic of special concern for airline companies, especially with the rapid growth of civil aviation transportation market demand and the intensified trend of competition among airline companies, airline companies are paying more and more attention to the key issue of market share. Therefore, this chapter focuses on the construction of a customer churn prediction model and the analysis of important factors affecting customer churn to provide a theoretical basis and scientific reference for airline companies to maintain their market share and competitive advantage. The key conclusions are as follows.

6.1. Dominance of hedonic services in mitigating customer churn

Hedonic service attributes, particularly inflight entertainment (importance score: 0.13), seat comfort (0.09), and type of travel (0.08)-exhibit the strongest influence on reducing customer churn. The nonlinear relationship between hedonic service quality and churn highlights a threshold effect: improvements only significantly reduce attrition when service scores exceed 3 (on a 5-point scale). At scores above 3, customers transition from "lost/remedial" to "positive/potential" states, with near-zero churn observed at the highest quality level (score = 5). This underscores that hedonic services act as differentiators that exceed baseline expectations, fostering emotional engagement and loyalty. For instance, business travelers who prioritize comfort and entertainment during frequent flights show markedly reduced churn when hedonic services surpass this critical threshold.

6.2. Paradoxical role of utilitarian services

While utilitarian attributes (e.g., flight punctuality, online boarding) are traditionally viewed as foundational, their impact on churn is limited or even counterproductive. Notably:

Negative marginal utility: Improvements in "convenient departure/arrival times" inadvertently increase churn by extending transfer waiting times, which amplifies dissatisfaction.

Diminishing returns: Utilitarian services like online support and gate location exhibit minimal importance scores, suggesting customers perceive them as basic entitlements rather than loyalty drivers.

This paradox implies that over-investment in utilitarian services may crowd out resources for hedonic enhancements while failing to address deeper emotional drivers of loyalty.

6.3. Nonlinear service expectations

The 3-point threshold (on a 5-point scale) emerges as a critical benchmark:

Below 3: Customers perceive services as failing to meet minimum expectations, entering a "lost/remedial" state where churn is inevitable.

Above 3: Incremental improvements yield disproportionate reductions in churn, validating the "hygiene-motivator" theory (Herzberg, 1959), where hedonic services act as motivators that transform satisfaction into loyalty.

The findings from this study provide several avenues for further investigation. First, Future research could explore how the effects of hedonic and utilitarian service quality evolve over time. Tracking changes in customer satisfaction and churn across multiple service improvements would help to identify long-term patterns and more accurately measure the impact of service enhancements. Second, Given the differential effects of hedonic and utilitarian factors across various customer segments (e.g., business vs. personal travelers), research could investigate how personalized service offerings (based on customer preferences, demographics, or travel type) might further enhance customer retention. This could include testing specific hedonic service enhancements targeted at business travelers or more utilitarian-focused improvements for leisure travelers. Finally, it would be also beneficial to compare the findings in the airline industry to other service-oriented industries, such as hospitality or transportation, to determine whether similar patterns emerge across different service contexts. This could shed light on whether hedonic and utilitarian factors have universal or industry-specific effects on customer loyalty and churn.

CRediT authorship contribution statement

Qiang Li: Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Funding acquisition, Formal analysis, Data curation. **Yuangang Li:** Writing – original draft, Validation, Supervision, Formal analysis, Data curation. **Ranzhe Jing:** Writing – review & editing, Validation, Methodology.

Acknowledgements

This work was supported by Shanghai Science and Technology Plan Project (No. 23692106100, 23DZ2204500). This work was supported by The Statistical Science Research Project of China (No. 2023LY035) and the China Postdoctoral Science Foundation (No. 2024M761198).

Data availability

The data that has been used is confidential.

References

Agarwal, I., Gowda, K.R., 2021. The effect of airline service quality on customer satisfaction and loyalty in India. Mater. Today Proc. 37, 1341–1348.
An, M., Noh, Y., 2009. Airline customer satisfaction and loyalty: impact of in-flight service quality. Service Business 3 (3), 293–307.
Athanassopoulos, A.D., 2000. Customer satisfaction cues to support market segmentation and explain switching behavior. Bus Res 47 (3), 191–207.
Belobaba, P., Odoni, A., Barnhart, C., 2015. The Global Airline Industry, second ed. John Wiley & Sons.
Blodgett, J.G., Li, H., 2007. Assessing the effects of post-purchase dissatisfaction and complaining behavior on profitability: a Monte Carlo simulation. J. Consumer Satisfaction, Dissatisfaction Complain. Beyond Behav. 20, 1–14.
Breiman, L., 1996. Bagging predictors. Mach. Learn. 24 (2), 123–140.
Calisir, N., Basak, E., Calisir, F., 2016. Key drivers of passenger loyalty: a case of Frankfurt-Istanbul flights. J. Air Transport. Manag. 53, 211–217.
Chen, P.T., Hu, H.H.S., 2013. The mediating role of relational benefit between service quality and customer loyalty in airline industry. Total Qual. Manag. Bus. Excel. 24 (9–10), 1084–1095.
De Caigny, A., Coussement, K., De Bock, K.W., Lessmann, S., 2020. Incorporating textual information in customer churn prediction models based on a convolutional neural network. Int. J. Forecast. 36 (4), 1563–1578.
DeMaris, A., 1995. A tutorial in logistic regression. J. Marriage Fam. 956–968.
Forgas, S., Moliner, M.A., Sanchez, J., Palau, R., 2010. Antecedents of airline passenger loyalty: low-cost versus traditional airlines. J. Air Transport. Manag. 16 (4), 229–233.
Gao, K., Yang, Y., Qu, X., 2021. Examining nonlinear and interaction effects of multiple determinants on airline travel satisfaction. Transport. Res. Transport Environ. 97, 102957.
Gordini, N., Veglio, V., 2017. Customers churn prediction and marketing retention strategies. An application of support vector machines based on the AUC parameter-selection technique in B2B e-commerce industry. Ind. Mark. Manag. 62, 100–107.
He, B., Shi, Y., Wan, Q., et al., 2014. Prediction of customer attrition of commercial banks based on SVM model. Procedia Comput. Sci. 31, 423–430.
Huang, B., Kechadi, M.T., Buckley, B., 2012. Customer churn prediction in telecommunications. Expert Syst. Appl. 39 (1), 1414–1425.
Hussain, R., Al Nasser, A., Hussain, Y.K., 2015. Service quality and customer satisfaction of a UAE-based airline: an empirical investigation. J. Air Transport. Manag. 42, 167–175.
Kamakura, W., Mela, C., Ansari, A., Bodapati, A., Fader, P., Iyengar, R., et al., 2005. Choice models and customer relationship management. Mark. Lett. 16 (3), 279–291.
Klein, T., 2015. U.S. Airline passenger satisfaction survey'. https://www.kaggle.com/teejmahal20/airline-passenger-satisfaction.
Kramer, O., 2013. K-nearest neighbors. In: Dimensionality Reduction with Unsupervised Nearest Neighbors. Springer, Berlin, Heidelberg, pp. 13–23.
Noble, W.S., 2006. What is a support vector machine? Nat. Biotechnol. 24 (12), 1565–1567.
Noviantoro, T., Huang, J.P., 2022. Investigating airline passenger satisfaction: data mining method. Research in Transportation Business & Management 43, 100726.
Shahin, A., Zairi, M., 2009. Kano model: a dynamic approach for classifying and prioritising requirements of airline travellers with three case studies on international airlines. Total Qual. Manag. Bus. Excel. 20 (9), 1003–1028.
Sharma, A., Panigrahi, P.K., 2011. A neural network based approach for predicting customer churn in cellular network services. Int. J. Comput. Appl. Technol. 27 (11), 26–31.
Tsai, C.F., Lu, Y.H., 2009. Customer churn prediction by hybrid neural networks. Expert Syst. Appl. 36 (10), 12547–12553.
Vafeiadis, T., Diamantaras, K.I., Sarigiannidis, G., Chatzisavvas, K.C., 2015. A comparison of machine learning techniques for customer churn prediction. Simulat. Model. Pract. Theor. 55, 1–9.