aerospace

Article

Prediction of Willingness to Pay for Airline Seat Selection Based on Improved Ensemble Learning

Zehong Wang [1], Xiaolong Han [1], Yanru Chen [2], Xiaotong Ye [1], Keli Hu [1] and Donghua Yu [1,*]

[1] Department of Computer Science and Engineering, Shaoxing University, Shaoxing 312000, China; 18164319@usx.edu.cn (Z.W.); 18145304@usx.edu.cn (X.H.); csyxt@usx.edu.cn (X.Y.); hukeli@usx.edu.cn (K.H.)
[2] Department of Music, Shaoxing University, Shaoxing 312000, China; 18054107@usx.edu.cn
* Correspondence: donghuayu163@163.com

Abstract: Airlines have launched various ancillary services to meet their passengers' requirements and to increase their revenue. Ancillary revenue from seat selection is an important source of revenue for airlines and is a common type of advertisement. However, advertisements are generally delivered to all customers, including a significant proportion of people who do not wish to pay for seat selection. Random advertisements may thus decrease the amount of profit generated since users will tire of useless advertising, leading to a decrease in user stickiness. To solve this problem, we propose a Bagging in Certain Ratio Light Gradient Boosting Machine (BCR-LightGBM) to predict the willingness of passengers to pay to choose their seats. The experimental results show that the proposed model outperforms all 12 comparison models in terms of the area under the receiver operating characteristic curve (ROC-AUC) and F1-score. Furthermore, we studied two typical samples to demonstrate the decision-making phase of a decision tree in BCR-LightGBM and applied the Shapley additive explanation (SHAP) model to analyse the important influencing factors to further enhance the interpretability. We conclude that the customer's values, the ticket fare, and the length of the trip are three factors that airlines should consider in their seat selection service.

Keywords: ensemble learning; ancillary service; LightGBM

check for updates

Citation: Wang, Z.; Han, X.; Chen, Y.; Ye, X.; Hu, K.; Yu, D. Prediction of Willingness to Pay for Airline Seat Selection Based on Improved Ensemble Learning. *Aerospace* **2022**, *9*, 47. https://doi.org/10.3390/aerospace9020047

Academic Editor: Cheng-Lung (Richard) Wu

Received: 29 November 2021
Accepted: 14 January 2022
Published: 19 January 2022

1. Introduction

With the development of airline business, ancillary services [1–7] that satisfy passengers' personal requirement are becoming increasingly important for airlines. Ancillary revenue has already played a vital role in airline profit and greatly increases the amount of extra financial revenue for airlines. By improving the quality of ancillary services, airlines increase their user satisfaction [8,9] and the adhesiveness of customers [2,7], which enhances their competitiveness and prevents homogeneity. Due to the worldwide spread of COVID-19, the global market for airlines has reduced dramatically [10–17]. Airline companies are, thus, urgently seeking extra profit to reduce fiscal pressure, leading to more serious competition based on ancillary services.

Airline ancillary revenue refers to income beyond the ticket fare and acts as a directly recommended service or implicit travel experience. Ancillary services are rapidly growing due to the fast-growing airline market (2007∼2018) and the impact of COVID-19 (2019∼2021). Ancillary revenue [18] greatly increased from \$2.1 billion to \$35.2 billion for the top 10 airlines within 12 years (2007∼2018). The significant growth in airline business in these years brings a great potential market for ancillary service.

Owing to the impact of the pandemic, the airline market faced a dramatic regression (2019∼2021), compelling airlines to seek revenue other than from flight tickets [12,14]. Therefore, establishing ancillary services is significantly important for airlines due to the ability to increase the airline's revenue. This also serves as an approach to solving the problem of customer churn and ensuring the resources are adequately utilized.

According to [19], the airline ancillary service is divided into five categories: (1) a la carte features, (2) commission-based products, (3) frequent flier activities, (4) advertising sold by the airline, and (5) the a la carte components associated with a fare or product bundle. From all of them, a la carte is the most general service that increases the revenue of airlines dramatically. A la carte features consist of multiple services that improve the travel experience, including onboard sales, extra baggage allowance, onboard Wi-Fi, and seat selection.

Seat selection [2,5,20,21] is one of the most common ancillary services chosen by passengers. This service refers to passengers choosing their preferred seats and paying for them willingly. For example, if passengers want to sit in the first row of the economy class for more legroom, they can spend extra money to reserve those seats in advance. Many reputable airlines around the world already provide this service and obtain revenue based on this service.

However, it is difficult for airlines to identify which passengers are willing to pay for seat selection since passengers who choose this service make up only a small part of all passengers. If the service is recommended to all customers, not only will advertising resources be wasted but passengers may also tire of useless advertisements, leading to a negative impact on customer satisfaction. Thus, how to precisely predict passengers' willingness to pay for seat selection must be urgently solved for airlines to save in advertisement resources and to increase their ancillary profits.

In this paper, we apply the air passenger seat selection dataset provided by Neusoft as the research object to analyse important factors in the willingness to pay for a seat selection service. In particular, we propose a machine-learning-based model named Bagging in Certain Ratio Light Gradient Boosting Machine (BCR-LightGBM) to predict the willingness of passengers to pay for such services. We conduct extensive experiments to demonstrate the effectiveness of BCR-LightGBM, showing the ability to capture rules between features. Then, we visual a decision tree in BCR-LightGBM and study two typical samples based on the visualization. To further enhance the interpretability, we use a Shapley additive explanation (SHAP) [22] model to analyse the feature importance and give our recommendations. Our contributions are summarized as follows:

1. We study the seat selection service from the perspective of passengers' willingness to pay, create new features from the original dataset, and propose an ensemble model, named BCR-LightGBM, to predict the willingness of passengers to pay for seat selection.
2. The experimental results show that BCR-LightGBM outperformed all 12 comparison models in terms of the AUC and F1-score.
3. We demonstrate the rules learned by BCR-LightGBM by visualizing the decision-making phase of two typical samples and analyse the important factors based on the SHAP model.

2. Related Work

2.1. Airline Ancillary Service

As the representative of high-level transportation, airlines are expected to acquire extra revenue from ancillary services and to satisfy passengers' personal requirements. How to increase their ancillary revenue has become a research hotspot for airlines. Chen et al. [23] studied passenger value on the air market between the Taiwan region and China's mainland. The results demonstrated that business travellers are less likely to perceive a trade-off compared with non-business travellers. Correia et al. [24] studied customers' preferences for ancillary services provided by low-cost airlines.

The research found that low-cost passengers are sensitive to the price of ancillary services. O'Connell et al. [1] employed an online survey to research the preference of travellers to ancillary service. They found that airport car parking and checked baggage charges were the most acceptable services. Wittmer et al. [20] studied the customer value and ancillary services based on a European network carrier's economy class.

The research revealed that the key point of passengers' willingness to pay for ancillary services is their perception of the importance of these services. Warnock-Smith et al. [2] investigated the relation between the willingness of passengers to pay for ancillary services and the pricing of the service. The work found that passengers prefer to choose necessary services that enhance the travel experience, e.g., seat selection, instead of optional services.

These studies examined ancillary services overall instead of focusing on a particular service. Han et al. [25] analysed the role of in-flight food and beverage in re-flying intention. Specifically, the quality of food and beverage, the reasonableness of the price, the airline image, and satisfaction were positive factors that influence re-flying intention. Klislinar et al. [4] analysed the relation between four main factors and the revenue generated from ancillary services based on a survey for Garuda Indonesia customers.

The results showed that passengers valued unbundled products more. Chiambaretto et al. [5] researched the willingness to choose ancillary services for air passengers on long-haul airlines. Five ancillary services, i.e., checked baggage, in-flight meal, seat selection, priority boarding, and onboard Wi-Fi, were analysed and revealed that leisure passengers were more likely to pay for extra services.

Influenced by the COVID-19 pandemic, the strategies for ancillary services by airlines have greatly changed. Various dynamic pricing strategies on airline ancillary services have been proposed to further increase the amount of extra revenue, mitigating the shortage of funds [15]. Vinod et al. [12] proposed an airline revenue planning method, including scheduling, airline pricing, and revenue management, to mitigate the volatility of airline revenue due to COVID-19.

Shukla et al. [26] proposed a dynamic, customer-specific pricing recommendation framework to increase the revenue of airline ancillary services. Compared with human rule-based approaches, the framework dramatically improved the expected revenue in online testing, showing the great potential of machine learning in decision-making.

Kolbeinsson et al. [7] proposed a dynamic and personalized pricing strategy based on flight characteristics and customer needs. The system greatly surpassed human-curated rules over a six-month live-implementation testing. Zhao et al. [6] analysed the passenger's willingness to pay for ancillary services through pricing strategies. By analysing the relationship between the pricing of services and the willingness to pay, that paper further proposed a dynamic pricing model for ancillary services to increase the extra revenue. Moreover, Shaw et al. [27] studied how to increase revenue from third-party ancillary services further increasing the number of sources of ancillary revenue.

2.2. Seat Selection

Profit from payable seat selection occupies a great proportion of ancillary profit. Rouncivell et al. [3] utilized UK domestic flights to study the willingness to pay for airline seat selection. They found that ticket fare was an important factor for both business and non-business travel and that passengers who chose the service in the past were more likely to choose it again.

Shao et al. [28] analysed five intercontinental routes from major European airlines to propose a statistical model for advanced seat reservation. The results showed that passengers generally avoid middle seats and prefer to sit in the first row, which provided an empirical foundation for seat selection services.

Zhou et al. [29] focused on the Chinese market to analyse the influencing factors for seat selection in economy class. They concluded that the length of the trip, the seat comfort and convenience, and payment and consumption situations greatly influenced the willingness of passengers to pay.

Yoon et al. [30] focused on customers' demands being uncertain and analysed how to maximize airline revenue by providing payable upgrade options, especially for seat assignment problems. This work analysed the willingness of passengers to pay and proposed some suggestions for airlines. However, to the best of our knowledge, no research has focused on the prediction of passengers' willingness to pay for seat selection.

If passengers who are willing to pay cannot be precisely predicted and targeted by advertisements, airline profits may decrease since customers may tire of random advertising. To solve this problem, in this paper, we propose a model to predict passengers who are willing to pay for seat selection and provide corresponding recommended services to them.

2.3. LightGBM

LightGBM (Light Gradient Boosting Machine) [31] is an improved Gradient Boosting Decision Tree (GBDT) [32] combining the decision tree and boosting methods. The essence of GBDT is to take the value of the negative gradient of the loss function in the current model as the approximation of the residual and to iteratively train multiple decision trees according to that value. However, the traditional GBDT model has some shortcomings, e.g., difficult parallelization, high computational cost, and not being suitable for high-dimensional sparse data.

LightGBM overcomes the shortcomings of traditional GBDT by supporting parallelized training to achieve fast speeds and low memory consumption when processing huge amounts of data. The biggest difference between LightGBM and other GBDT models is that the other models pre-sort the feature values and find the optimal division point according to the sorting result. The implementation is simple but difficult to be optimized. When the dataset is large, the training process occupies a great deal of memory, which leads to a waste of CPU cycles and reduces the training speed. To solve this problem, LightGBM applies a histogram-based algorithm to discretize the numerical features into K discrete values and to pick the value that achieves the highest accumulated number as the split point.

LightGBM utilizes a sampling algorithm named Gradient-based One-side Sampling (GOSS) to reduce the number of instances. The algorithm excludes most low-gradient samples and calculates the information gained by the other samples, maintaining the performance of the model when the training dataset is reduced. To further improve the computational speed, LightGBM applies a bundle method, Exclusive Feature Bundling (EFB), to combine features that have small conflicts or are totally exclusive to reduce the number of features. Although the algorithms mentioned above greatly reduce the computational consumption, the performance of the model is also decreased. To handle this issue, LightGBM introduces a leaf-wise splitting mechanism, which effectively reduces the loss and increases the precision.

3. Methodology

How to predict the willingness to pay for airline seat selection is an urgent problem that needs to be solved. We utilize real air passenger history provided by Neusoft (described in Section 4.1) to predict their willingness since the corresponding dataset is hard to collect from individuals rather than from airlines. First, we construct and select new features to overcome the data sparsity and the curse of dimensionality. Then, we propose Bagging in Certain Ratio LightGBM (BCR-LightGBM) to solve the issue of imbalance.

3.1. Feature Construction

To overcome the problem of data sparsity, we construct new features on the basis of the original dataset. There are three types of data in the dataset, i.e., date, numeral, and category; we apply different transformations on each. For the date and time, e.g., “16 December 2018 20:00”, we construct two features to indicate the season (month-wise) and time period (hour-wise), shown as follows:

$$seg\_dep\_month = \begin{cases} a, & x \in \{1,2,12\} \\ b, & x \in \{3,4,5\} \\ c, & x \in \{6,7,8\} \\ d, & x \in \{9,10,11\} \end{cases} \tag{1}$$

where x is the month of the flight.

$$seg\_dep\_hour = \begin{cases} a, & x \in \{6,7,8,9,10,11\} \\ b, & x \in \{12,13,14,15,16,17\} \\ c, & x \in \{18,19,20,21,22,23\} \\ d, & x \in \{24,0,1,2,3,4,5\} \end{cases} \tag{2}$$

where x is the hour of the flight.

For numerical and categorical features, the names of the features are divided into two parts, i.e., characteristic (prefix) and time interval (suffix). The characteristic denotes the history of the passenger or the inherent property of the flight. For instance, "dist_all_cnt" indicates the total mileage of a passenger and "pax_fcny" indicates the fare of the flight ticket. The time interval denotes the time scope of the characteristic, e.g., "dist_all_cnt_m3" represents the total mileage of a specific passenger collected from three months ago to the current time. The time interval includes five scopes, i.e., 3 months, 6 months, 1 year, 2 years, and 3 years. For simplicity, we call the characteristics prefixes and time intervals suffixes in the following.

We observe that the issue of sparsity is severe for both numerical and categorical features. To overcome the sparsity for the numerical features, we directly conduct statistics-based transformation on the original numerical features, i.e., maximum, minimum, mean, and variance. On the basis of the transformation, we improve the interpretability of each numerical feature. The features newly formed are named "prefix_max", "prefix_min", "prefix_mean", and "prefix_std" for each prefix.

$$prefix\_max = max(features\_with\_same\_prefix) \tag{3}$$
$$prefix\_min = min(features\_with\_same\_prefix) \tag{4}$$
$$prefix\_mean = mean(features\_with\_same\_prefix) \tag{5}$$
$$prefix\_std = std(features\_with\_same\_prefix) \tag{6}$$

For each categorical feature, we define two sub-flags, named secondary indexes, to indicate the relationship between the target and the feature. The secondary indexes are represented as "prefix_T" and "prefix_F", where "T" denotes a passenger paying for seat selection services and "F" is a passenger who does not. The constructional rule is shown as follows:

1. According to the target, all values are divided into two sets for each categorical feature. These two sets are denoted as S_0 and S_1, which contain values $\{x|x \in \{0,1\}\}$.
2. If S_0 or S_1 contains 0, delete it from the set.
3. Construct two new features, "prefix_T" and "prefix_F", based on the transformation rule followed:

$$prefix\_T = \begin{cases} T, & \exists x \in S_1 \\ U, & \forall x = 0 \\ F, & others \end{cases} \tag{7}$$

$$prefix\_F = \begin{cases} T, & \exists x \in S_0 \\ U, & \forall x = 0 \\ F, & others \end{cases} \tag{8}$$

where x represents the values of a sub-label, and $prefix$ is the prefix of the feature.

3.2. Feature Selection

In addition to sparsity, the dataset suffers from dimensionality. In this section, we apply Pearson correlation coefficient-based [33] and chi-square test-based [34] feature selection techniques to reduce the dimension of numerical and categorical features, respectively. To select the numerical features, we perform the process below:

1. Calculate the Pearson correlation coefficient between any two features and sort them in descending order.

2. Set a preserve threshold and delete the threshold to indicate the state of features.
3. For each feature set (a, b), if the correlation coefficient is less than the preserve threshold, we set a as a preserve state; if the correlation coefficient is greater than the delete threshold, we set a as a delete state.
4. If feature a is in the delete state, delete it unless b is in the delete state.
5. If feature a is in the preserve state, delete feature b unless b is in the preserve state.
6. If features a and b are in neither the delete state nor the preserve state, delete the feature with the smaller variance.

For each categorical feature, we validate the mutual independence between it and the target through a chi-square test. If the feature is independent from the target, we directly delete the feature.

3.3. BCR-LightGBM

Predicting whether air passengers are willing to pay for seat selection is a binary task. In this section, we illustrate the structure of Bagging in Certain Ratio LightGBM (BCR-LightGBM). To ensure the robustness, multiple LightGBMs are assembled through a bagging method [35]. Bagging ensembles are multiple models that were trained by subsets extracted from the original dataset through bootstrap sampling. The result is obtained by applying an average or voting strategy. By utilizing bagging, the effectiveness and stability are improved and the variety of the model is lowered. However, the bootstrap sampling does not change the data distribution; thus, it cannot overcome the data imbalance in the original dataset.

To mitigate the imbalance of the original dataset, we only sample the negative samples in the sampling phase and then combine them with the positive samples to create the subset. Note that the ratio between positive and negative needs to be pre-assigned since different ratios lead to different results. Then, each LightGBM is trained through the subsets sampled, and the prediction is the average of their results. The training process of BCR-LightGBM is shown in Algorithm 1.

Algorithm 1: The training process of BCR-LightGBM.

```
Input  :Data, $Num_{subset}$, Ratio
Output:BCR-LightGBM classifier
Step 1: Clean the original dataset
Data = Remove_null(Data) # Remove all samples with null value
Data = Translate(Data, 0) # Translate all values above 0
Step 2: Construct new features
Data = Transform(Data, Datetime) # Construct date time features
Data = Transform(Data, Int, Float) # Construct numerical features
Data = Transform(Data, Object) # Construct categorical features
Step 3: Select features
Data = Select(Data, Pearson) # Select numerical features
Data = Select(Data, Chi-square) # Select categorical features
Step 4: Split dataset
$Data_{pos}, Data_{neg} \leftarrow$ Data.split() # Split the dataset into postive and negative
Step 5: Create subsets
for $i \leftarrow 1$ to $Num_{subset}$ do
    $Num_{pos} \leftarrow$ Num($Data_{pos}$) # Count the number of positive instances
    $Sampled_{neg} \leftarrow$ Sampling($Data_{neg}, Ratio \times Num_{pos}$) # Sample negative instances
    $Subset_i \leftarrow$ Combine($Data_{pos}, Sampled_{neg}$) # Create subset
end
Step 6: Train LightGBM
for $i \leftarrow 1$ to $Num_{subset}$ do
    $clf_i \leftarrow$ LightGBM.train($Subset_i$) # Train multiple LightGBMs through subsets
end
Step 7: Ensemble the prediction
BCR-LightGBM $\leftarrow$ Bagging($\{clf_i\}_{i=0}^{n}$) # Ensemble LightGBMs throught Bagging
```

4. Experimental Results

In this section, we compare the proposed BCR-LightGBM against various machine learning algorithms and sampling-based methods. Then, we illustrate the decision-making procedure of a decision tree in BCR-LightGBM on two real samples to demonstrate the learned mode of the model. Furthermore, we analyse the feature importance through a SHAP model to improve the interpretability. Extensive experiments are conducted on a 64-bit Ubuntu 16.04 operating system. The setting environment is as follows: CPU: Intel (R) Xeon(R) Silver 4114 CPU @ 2.20 GHz, memeory: 64 RAM, and graphics: GeForce GTX 1080 Ti.

4.1. Data Description

This paper uses the dataset of air passenger willingness to pay for seat selection provided by Neusoft http://fwwb.org.cn/attached/file/20201211/20201211132638_47.zip (accessed on 29 November 2021), consisting of flight information, passenger history, and customer characteristics, which are shown in Table 1. The dataset comprises 23,432 samples, and the feature dimension is 657, which increases the risk of dimensionality [36]. Note that the dataset contains features with the same prefix.

For example, there are five features prefixed with “cabin_f_cnt”, i.e., “cabin_f_cnt_m3”, “cabin_f_cnt_m6”, “cabin_f_cnt_y1”, “cabin_f_cnt_y2”, and “cabin_f_c-nt_y3”, which represents the number of people who took first-class flights in the last x months/years, where $m3$, $m6$, $y1$, $y2$, and $y3$ represents 3 months, 6 months, 1 year, 2 years, and 3 years, respectively. Moreover, the dataset is especially sparse, where nearly 70% of the values are zero.

Positive samples in the dataset indicate the people who paid for seat selection, and negative samples represent the people who did not. Note that the dataset is extremely imbalanced [37,38], where the ratio between positive and negative is $1:15$. The situation is common for ancillary services since the majority of people do not choose extra services even though ancillary profits dramatically aid airlines.

Table 1. Description of the dataset provided by Neusoft.

Description	Samples	Positive	Negative	Features	Datatime	Float	Int	Object
Number	23,432	1475	21,952	657	1	90	444	123

4.2. Metrics

In this section, we introduce the five metrics, i.e., Accuracy, Precision, Recall, F1-score, and ROC-AUC, used to evaluate the performance of the model. Accuracy is the simplest metric, which is defined as the number of correct predictions divided by the total number of predictions, indicating the proportion of correct predictions. Precision and Recall are two mutually influencing indicators, where Precision indicates the correctness of the prediction and Recall indicates the prediction performance for users who are willing to pay for seats. However, there are many cases in which these metrics are not good enough to indicate of the model performance.

A scenario is when the class distribution is imbalanced, e.g., the case in the experiment. In this case, even if the model predicts all samples as the most frequent class, the performance of these metrics would obtain a high accuracy rate. However, the model is not learning anything and is simply predicting every sample as the top class. For the dataset used in the experiment where the negative class occupies around 93.7% samples, if the model predicts all instances as negative, it would result in a 93.7% accuracy.

To cover the shortage of these metrics and better indicate the model performance, we further introduce the F1-score and ROC-AUC (area under the receiver operating characteristic curve). The F1-score combines Precision and Recall into a single metric, with the cases in which both Precision and Recall are important. The indicator is the harmonic average of Precision and Recall, and always achieves a trade-off between them, which is generally applied to indicate the overall performance of model when the dataset is imbalanced.

ROC-AUC indicates the area under the ROC (receiver operating characteristic curve) where the ROC is used to show the performance of a binary classifier. Specifically, the ROC-AUC is an aggregated measure of performance of a binary classifier on all possible threshold values. Thus, the indicator is not sensitive to threshold. When the ratio between positive samples and negative samples changes, the ROC-AUC value does not change dramatically. *TP* is the number of instances correctly classified as positive, *TN* is the number of instances correctly classified as negative, *FP* is the number of instances incorrectly classified as positive, and *FN* is the number of instances incorrectly classified as negative.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{12}$$

$$AUC = \int_0^1 \frac{TP}{TP + FN} d\left(\frac{FP}{TN + FP}\right) \tag{13}$$

4.3. Comparison Models

In this section, we introduce various comparison models used in the experiments, including machine-learning methods and sampling-based methods.

1. LR (Logistic Regression) is a simple linear model that can be easily interpreted, where the performance greatly relies on feature engineering.
2. KNN (K-Nearest Neighbours) is a learning-free model that classifies a sample based on the k-nearest samples in the feature space.
3. SVM (Support Vector Machine) is not sensitive to outliers due to the inherent properties of support vectors. However, the kernal function should be dedicated and designed to fit the input space.
4. AdaBoost (Adaptive Boosting) is a boosting method, which dynamically adjuncts the weight of each base learner to improve the robustness.
5. GBC (Gradient Boosting) is a boosting method in which the objective is to find the optimal solution in the parameter space by fitting the residual error of a previous learner.
6. RF (Random Forest) adopts bagging to improve the robustness, where the decision tree is a base learner that has been widely used in various fields.
7. XGBoost (eXtreme Gradient Boosting) [39] is an extension of GBC that achieves better performance and scalability.
8. LightGBM (Light Gradient Boosting Machine) [31] is an extension of GBC. Compared with XGBoost, LightGBM is faster and lighter. Note that LightGBM is the base learner in BCR-LightGBM.
9. RUS (Random Under Sampling) randomly samples negative instances until the number is the same as that of positive instances.
10. ROS (Random Over Sampling) randomly samples positive instances until the number is same as that of negative instances.
11. SMOTE (Synthetic Minority Over-Sampling) [40] is an over-sampling method, creating synthetic instances for minorities based on the nearest neighbours.
12. SMOTE-ENN (Synthetic Minority Over-Sampling and Edited Nearest Neighbours) [41] is the combination of SMOTE and ENN, which applies ENN to clean the samples created by SMOTE.

4.4. Comparative Analysis

To demonstrate the superiority of BCR-LightGBM, we compare the performances of the model against existing machine-learning methods and sampling methods in this section. For the proposed BCR-LightGBM, we set the ratio between positive samples and negative samples to 1:3.

Table 2 shows the performance comparison between the proposed BCR-LightGBM and machine-learning models without sampling. Note that BCR-LightGBM outperforms all methods in terms of the F1-score and AUC, which are two widely used indicators when the dataset is imbalanced. Furthermore, BCR-LightGBM achieves the narrowest gap between Precision and Recall. For other compared methods, the Precision is much higher than the Recall, indicating that these models only find a small set of passengers who are willing to pay for seat selection when reducing the error rate.

In other words, these models only can identify instances of people who are the most likely to pay for seat selection. However, this limitation is unnecessary for airlines as the cost of advertising is not too unacceptable that messages cannot be advertised to a relatively large set of people. BCR-LightGBM achieves a desired Recall and an acceptable Precision, satisfying the requirements of airlines.

Table 3 shows the comparative results between the BCR-LightGBM and sampling-based methods. The proposed model achieves the best score in terms of Accuracy, Precision, F1-score, and AUC. Note that sampling-based methods are generally better than machine-learning models, which is reflected in the narrower gap between Precision and Recall. Furthermore, the performance of under-sampling-based methods (RUS and SMOTE-ENN) is worse than over-sampling-based methods (SMOTE and RUS) because under-sampling-based methods drop a large number of instances in the original dataset, leading to the model not effectively learning the characteristics of the discarded samples and increasing the possibility of under-fitting.

The over-sampling-based methods create new samples to mitigate the issue of imbalance, improving the performance even though noise is introduced. Although BCR-LightGBM applies an under-sampling method, its performance is better than that of over-sampling-based models since the ensemble strategy is utilized.

Table 2. Comparative analysis with machine-learning models. Due to the imbalance of the dataset, we mainly compare the performances of models in terms of the F1-score and AUC.

	Accuracy	**Precision**	**Recall**	**F1-Score**	**AUC**
LR	0.9371	0.5556	0.0034	0.0067	0.7170
KNN	0.9322	0.2738	0.0468	0.0799	0.6539
SVM	0.9370	0.3947	0.0102	0.0198	0.6394
AdaBoost	0.9366	0.3077	0.0054	0.0107	0.7070
GBC	0.9375	0.8235	0.0095	0.0188	0.7166
RF	0.9431	0.6972	0.1702	0.2736	0.7534
XGBoost	0.9374	0.6538	0.0115	0.0227	0.7455
LightGBM	0.9381	0.6444	0.0393	0.0741	0.7565
BCR-LightGBM	0.8926	0.2449	0.3390	**0.2843**	**0.7732**

Table 3. Comparative analysis with sampling-based methods. We note that the proposed BCR-LightGBM surpasses existing methods in terms of the Accuracy, Precision, F1-score, and AUC.

	Accuracy	**Precision**	**Recall**	**F1-Score**	**AUC**
RUS-LightGBM	0.6315	0.1125	0.7044	0.1940	0.7275
ROS-LightGBM	0.8079	0.1586	0.4766	0.2380	0.7430
SMOTE-LightGBM	0.7639	0.1372	0.5200	0.2171	0.7244
SMOTE-ENN-LightGBM	0.7226	0.1285	0.5892	0.2110	0.7357
BCR-LightGBM	**0.8926**	**0.2449**	0.3390	**0.2843**	**0.7732**

To further illustrate the performance of BCR-LightGBM, we plot the ROC (Receiver Operating curve) of machine-learning methods and sampling-based methods, as shown in Figure 1. In Figure 1a, we compare the proposed model against various machine-learning methods, and, in Figure 1b, we compare BCR-LightGBM with LightGBM based on different sampling strategies. The ROC demonstrates the trade-off between Precision and Recall. We note that the curve of BCR-LightGBM wraps around all other curves, indicating that the proposed model surpasses all other methods.

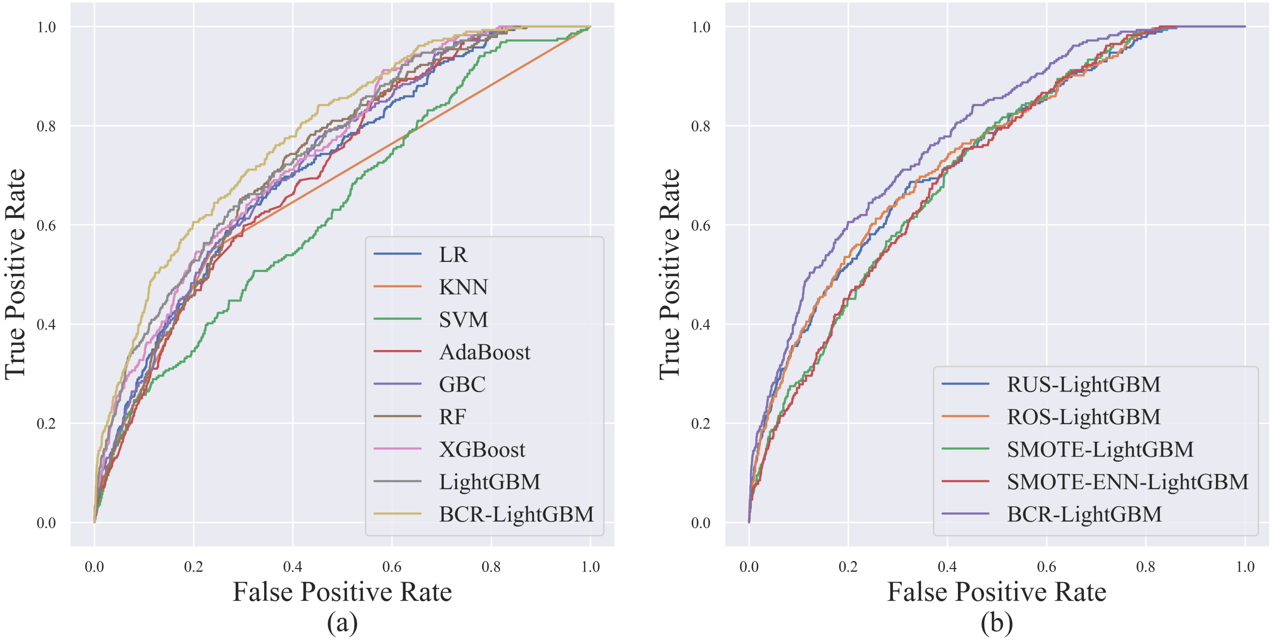

Figure 1. ROC curve compared with (**a**) machine-learning methods and (**b**) sampling-based methods. Note that BCR-LightGBM outperforms all baselines.

Note that the superiority of BCR-LightGBM is derived from the ability to correctly learn the relationship between important factors. The relation cannot be learned by other models. We attempted to analyse this from the perspective of model capability. A simple linear model, i.e., LR, cannot perform feature crossing, which limits its capability to learn the relation between features. KNN classifies samples through k-nearest neighbours in the original feature space, and the correlation between features cannot be identified.

Although SVM is not sensitive to outliers, it also cannot perform feature crossing and the kernel function needed to be dedicated in design. AdaBoost dynamically changes the weight of the base learner; however, the weight is sensitive to the data distribution. GBC finds the optimal solution through a descending gradient, which is dramatically influenced by an imbalance in the dataset. Although RF, XGBoost, and LightGBM achieve great performances in various fields, they are still weak when solving with data imbalances.

In summary, these models cannot solve or are weak when solving the issue of data imbalances, mainly in learning information from negative samples, which leads these models to not correctly find the relationship between important factors. However, the proposed BCR-LightGBM sets the ratio between the positive samples and negative samples at a certain value, mitigating the impact of negative samples and achieving the best performance.

For sampling-based methods, RUS causes information loss by dropping existing samples, and ROS magnifies the impact of outliers in positive samples. SMOTE creates new samples based on the samples in the dataset but introduces noise. Although SMOTE-ENN leverages ENN to clean the samples generated by SMOTE, the data distribution may be further misled due to the lack of prior data. To reduce the impact of noise and to avoid information loss, BCR-LightGBM applies random under sampling to avoid extra noise and uses an ensemble approach to learn all information in the dataset, thereby, improving the robustness.

4.5. Hyperparameter Analysis

To further analyse the performance of BCR-LightGBM, we conducted experiments to analyse two important hyperparameters, i.e., the ratio between negative samples and positive samples, and the number of LightGBMs in the model. Figure 2 shows the performance of BCR-LightGBM under different ratios between negative samples and positive samples. In the figure, α is the ratio between negative samples and positive samples. When $\alpha = 1$, the number of positive samples is equal to the number of negative samples. For a fair comparison, the number of LightGBMs is set at 100. Note that, if the ratio is too large, the model does not learn the correct relationship between features, thus, causing serious model degradation.

We observe that, when the ratio between negative samples and positive samples is 3:1, the model achieves the best performance in terms of the F1-score and AUC. In fact, with the increase in the ratio, the F1-score and AUC dramatically decreased and the gap between Precision and Recall becomes large. We do not use Accuracy to indicate the performance of the model because, even though all samples are predicted to be negative when the data distribution is imbalanced, the indicator still maintains a high level.

Figure 3 shows the impact of the number of LightGBMs (β) on the model performance. For simplicity, we select ROC-AUC as the indicator from the five matrices. To completely demonstrate the impact of the number of base classifiers, i.e., LightGBM, on the BCR-LightGBM, we conduct experiments when the ratio of negative samples and positive samples is in $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\}$.

From the figure, we observe that the increase in the number of LightGBMs can greatly improve the performance of the model since the model obtains a desirable performance gain even though the number of LightGBMs is relatively small. Moreover, with the increase in the number, the model shows strong robustness because the model rapidly converges even though slight fluctuation occurs. Note that, the performance of BCR-LightGBM generally converges when the number of LightGBMs is less than 50, thereby, demonstrating the robustness and stability of the model.

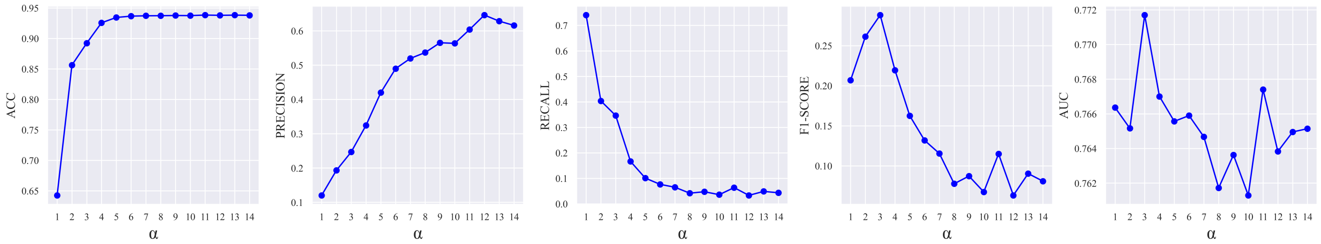

Figure 2. Performance analysis under different ratios between positive and negative samples where α indicates the ratio between negative samples and positive samples. Note that, when α is set at 3, BCR-LightGBM achieves the best performance in terms of the F1-score and AUC.

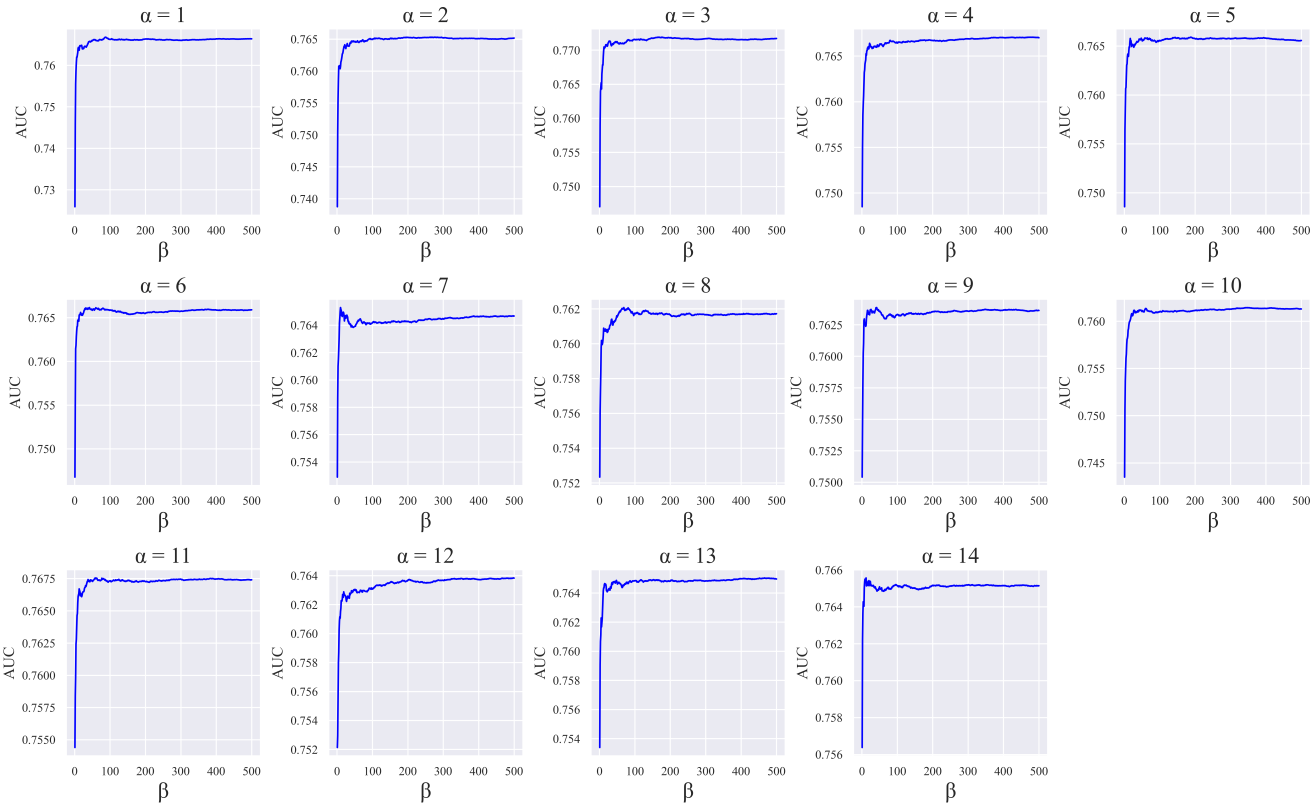

Figure 3. Performance analysis under a different number of base learners (LightGBM) in terms of the AUC. α indicates the ratio between negative samples and positive samples, and β denotes the number of LightGBMs.

4.6. Importance Analysis of Influencing Factors

In addition to comparing the performances between BCR-LightGBM and existing models, we also attempt to explain the rules learned by the model. Figure 4 illustrates a decision tree in the proposed model. For simplicity, we only visualize the top four layers. We set the layers of the model to seven to improve the capability. We note that the flight cabin ('seg_cabin') is located within the top of the tree, indicating that the cabin is most discriminative factor for seat selection in terms of the Gini index. Moreover, we find that some flight information, e.g., the tax of the ticket (pax_tax) and the month of travel (seg_dep_month), influences their willingness to pay.

In addition to that, the history of flights shows whether the passenger values also contribute to their willingness to pay, e.g., the number of paid seat selections (select_seat_cnt_max), number of seats by the window (select_window_cnt_var), number of international tickets (tkt_i_amt_max), point additions from airline mile accumulation (pit_add_air_cnt_y1), number of economy class flights (cabin_y_cnt_max), number of first-class flights (cabin_f_cnt_max), and member level (member_level).

To illustrate the rules learned by the model, we selected two typical samples from the dataset to simulate decision-making by the model, as shown in Table 4. The positive sample is a person who pays to select a seat, and the negative sample is one who does not. The features in the table match the corresponding split point in Figure 4. For the positive sample, we observe that the passenger is a frequent flyer since the value of corresponding factors is high, e.g., the times of economy class (cabin_y_cnt_var), the times of first class (cabin_f_cnt_max), and the total amount of international flight mileage accumulated (tkt_i_amt_max, tkt_i_amt_min).

The passenger always pays to select a seat (select_seat_cnt_max) and prefers to seat by the window (seat_window_cnt_var). Thus, the positive sample has a high customer value. We assume that the passenger generally takes business trips due to the frequency of flight and their willingness to pay for seat selection. Intuitively, a business traveller is generally willing to pay for seat selection to acquire seats that provide them with better rest. The model learns this mode following the orange arrow presented in Figure 4.

For the negative sample, we observe that the passenger does not always take flights due to the low number of flights in economy class (cabin_y_cnt_var) and in first class (cabin_f_cnt_max), the slow accumulation of points (pit_add_air_cnt_y1, pit_income_avg_amt_var), and the unwillingness to pay for seat selection (select_seat_cnt_max). Intuitively, these kinds of passengers are not willing to pay for seat selection. The model can learn this mode following the blue arrows in Figure 4.

Table 4. Two typical samples. The positive sample is a person who pays to select a seat. The negative sample is a person who does not pay to select a seat, representing people who do not always take flights. These two samples are used to demonstrate the learning mode of the decision tree in Figure 4.

Positive Sample				Negative Sample			
seg_cabin	1.0000	cabin_y_cnt_var	0.5983	seg_cabin	1.0000	cabin_y_cnt_var	0.0272
tkt_i_amt_max	0.5832	cabin_f_cnt_max	0.1921	tkt_i_amt_max	0.9280	cabin_f_cnt_max	0.0000
residence_country	0.0000	member_level	0.6000	residence_country	0.0000	member_level	0.0000
pit_add_air_cnt_y1	0.5930	seat_window_cnt_var	0.7583	pit_add_air_cnt_y1	0.0091	seat_window_cnt_var	0.0000
tkt_i_amt_min	0.0000	pit_income_avg_amt_var	0.6240	tkt_i_amt_min	0.0000	pit_income_avg_amt_var	0.0281
select_seat_cnt_max	0.6712	seg_dep_month	0.0000	select_seat_cnt_max	0.0000	seg_dep_month	0.2500
pref_orig_y3_T	0.0000	pax_tax	0.4031	pref_orig_y3_T	0.0000	pax_tax	0.1020
nation_name	0.0000	pref_orig_city_F	0.0000	nation_name	0.0000	pref_orig_city_F	0.0000

We further utilize a SHAP [22] model to enhance the interpretability of BCR-LightGBM. To interpret the results of ensemble models, SHAP [22] provides an approach to explain the prediction of ensemble models utilizing the contributions of allocation methods from cooperative games. The model considers the contribution of features for the prediction and calculates the feature importance based on that. Compared with feature importance derived from LightGBM to indicate the number of times to create a split point, the SHAP model explains the influences of each sample for the prediction and indicates the positive or negative effects on the prediction. To obtain the influence of each feature, SHAP calculates the Shapley value [42] of each feature.

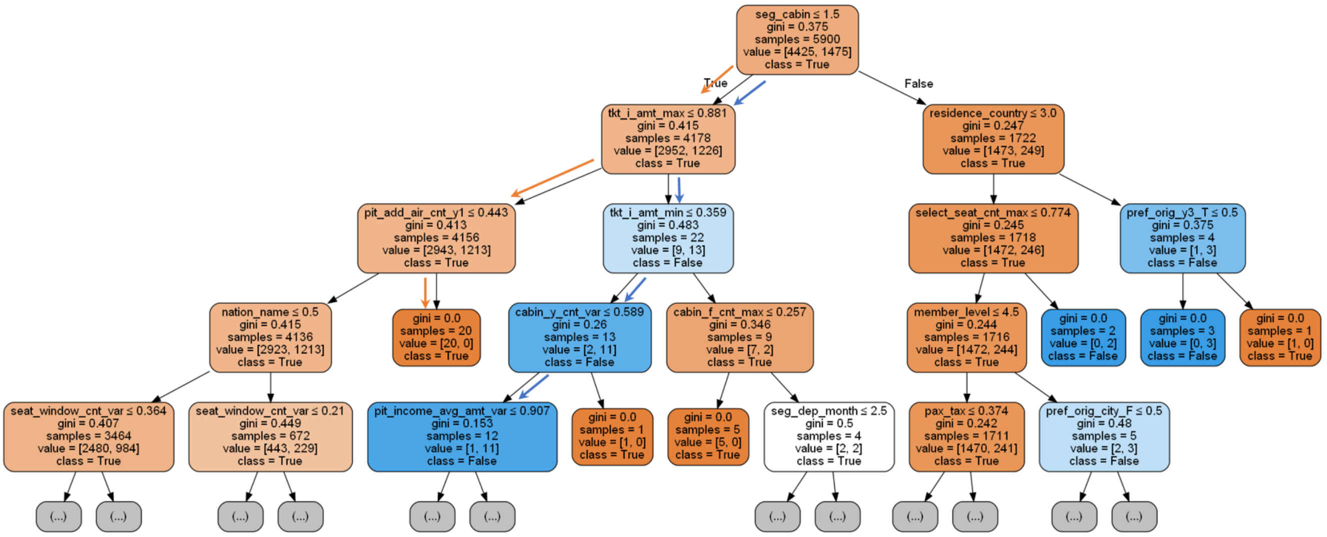
seg_cabin ≤ 1.5 gini = 0.375 samples = 5900 value = [4425, 1475] class = True
True
False
tkt_i_amt_max ≤ 0.881 gini = 0.415 samples = 4178 value = [2952, 1226] class = True
residence_country ≤ 3.0 gini = 0.247 samples = 1722 value = [1473, 249] class = True
pit_add_air_cnt_y1 ≤ 0.443 gini = 0.413 samples = 4156 value = [2943, 1213] class = True
tkt_i_amt_min ≤ 0.359 gini = 0.483 samples = 22 value = [9, 13] class = False
select_seat_cnt_max ≤ 0.774 gini = 0.245 samples = 1718 value = [1472, 246] class = True
pref_orig_y3_T ≤ 0.5 gini = 0.375 samples = 4 value = [1, 3] class = False
nation_name ≤ 0.5 gini = 0.415 samples = 4136 value = [2923, 1213] class = True
gini = 0.0 samples = 20 value = [20, 0] class = True
cabin_y_cnt_var ≤ 0.589 gini = 0.26 samples = 13 value = [2, 11] class = False
cabin_f_cnt_max ≤ 0.257 gini = 0.346 samples = 9 value = [7, 2] class = True
member_level ≤ 4.5 gini = 0.244 samples = 1716 value = [1472, 244] class = True
gini = 0.0 samples = 2 value = [0, 2] class = False
gini = 0.0 samples = 3 value = [0, 3] class = False
gini = 0.0 samples = 1 value = [1, 0] class = True
seat_window_cnt_var ≤ 0.364 gini = 0.407 samples = 3464 value = [2480, 984] class = True
seat_window_cnt_var ≤ 0.21 gini = 0.449 samples = 672 value = [443, 229] class = True
pit_income_avg_amt_var ≤ 0.907 gini = 0.153 samples = 12 value = [1, 11] class = False
gini = 0.0 samples = 1 value = [1, 0] class = True
gini = 0.0 samples = 5 value = [5, 0] class = True
seg_dep_month ≤ 2.5 gini = 0.5 samples = 4 value = [2, 2] class = True
pax_tax ≤ 0.374 gini = 0.242 samples = 1711 value = [1470, 241] class = True
pref_orig_city_F ≤ 0.5 gini = 0.48 samples = 5 value = [2, 3] class = False
(...)
(...)
(...)
(...)
(...)
(...)
(...)
(...)
(...)
(...)
(...)
(...)

Figure 4. Visualization of one decision tree in BCR-LightGBM. For simplicity, we only visualize the top four layers. The willingness to pay for seat selection is denoted by the colour, where orange is true, and blue is false. Note that the deeper the colour, the greater the likelihood.

Figure 5 shows the feature importance based on the Shapley value. Note that aircraft cabin (seg_cabin), ticket tax (pax_tax), ticket fare (pax_fany), the gap between current and recent travel date (recent_gap_day), and total international flight mileage (dist_i_cnt_max) have the greatest impacts on the willingness of passengers to pay for seat selection. According to the SHAP value, we note that flight information has a great influence on the prediction since five features are important, as presented in Figure 5, i.e., aircraft cabin (seg_cabin), ticket fare (pax_fany), ticket tax (pax_tax), and the month of flight (seg_deg_month).

We conclude that passengers who pay for better aircraft cabins with higher ticket fares and taxes are more likely to choose the extra seat selection service and that those who travel in fall or winter are more likely to pay for these services. Furthermore, passenger history, which denotes the customer value, also greatly influences the prediction result, i.e., the gap from the most recent flight (recent_gap_day), total mileage and international mileage (dist_all_cnt_mean, dist_i_cnt_max), and average ticket fare and international ticket fare (tkt_avg_amt_max, tkt_i_amt_max). In general, the higher the total mileage and international mileage, the higher the average and international ticket fare, and the more frequently a passenger travels, the more likely the passenger is to pay for a seat selection service.

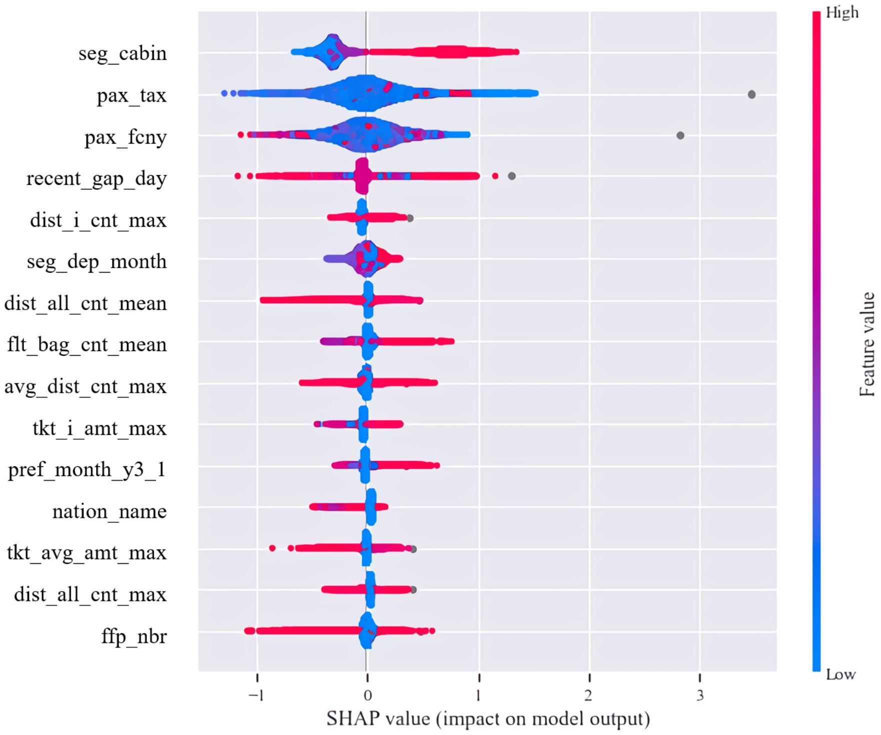

Figure 5. Feature importance analysis based on the SHAP value. For simplicity, we illustrate the top 15 most important features. If the SHAP value of a specific feature value is greater than 0, the feature value has a positive impact on the result and vice versa.

According to the sample analysis based on a visualization of the decision tree and SHAP-based feature importance analysis mentioned above, we conclude the following:

(1) Passengers who have a high customer value, reflected in the total fare of a ticket, the total mileage accumulated from flights, and the frequency of flights, will pay for seat selection. Thus, airlines can recommend seat selection services to them since they may pay more attention to their comfort on the plane.
(2) Passengers who choose airlines with higher ticket fares are more willing to pay for seat selection. The fare of the ticket also denotes their customer value when the customer history is difficult to acquire. Airlines can easily identify the willingness of

a passenger to pay for seat selection based on information from a single flight and can recommend seat selection services to them.

(3) Passengers who make international flights are more likely to pay for seat selection. We assume that this is because passengers like to have a more comfortable experience in long-haul flights. Therefore, airlines can recommend these services to passengers in long-haul flights.

5. Conclusions

Ancillary service revenue has become important for airlines in recent years. Under the impact of COVID-19, how to precisely provide personalized ancillary services to passengers to increase revenue and how to mitigate capital shortages are problems that need to be urgently solved. In this paper, we analysed seat selection services from the perspective of the prediction of the willingness of passengers to pay for seat selection and propose a machine-learning-based method to identify their willingness to pay for seat selection. Specifically, we proposed a model, named BCR-LightGBM, to identify passengers who are willing to pay for seat selection as the basis of recommendation.

We first preprocessed the original dataset to overcome the data sparsity and the curse of dimensionality inherent in the dataset. Then, the bagging method was applied, where positive samples and negative samples were combined at a specific ratio for multiple subsets to solve the problem of data imbalance. The experimental results demonstrated that the proposed model achieved 0.28 and 0.77 in terms of the F1-score and AUC, outperforming all existing machine-learning models and sampling-based methods.

Finally, we analysed two typical samples based on the visualization of a decision tree in BCR-LightGBM and applied a SHAP model to further enhance the interpretability by analysing feature importance. We note that customer value, ticket fare, and flight length had positive influences on the willingness to pay for seat selection. Based on this rule, airlines can recommend seat selection services to the corresponding passengers to increase their revenue.

The limitation of this research is that the number of samples is relatively small and cannot cover all situations regarding seat selection around the world. Thus, our conclusions may only be appropriate in similar cases to those contained in the dataset. In future research, we will collect more samples from different airlines to make the conclusions more convincing. We will further study the intrinsic properties of these important factors and mine knowledge from the dataset to guide the recommendation policies of airlines to increase revenue from other ancillary services, e.g., priority boarding, checked baggage, and onboard Wi-Fi.

Author Contributions: Conceptualization, Z.W. and D.Y.; methodology, Z.W. and D.Y.; software, Z.W. and X.H.; validation, Y.C. and K.H.; formal analysis, X.H.; investigation, Y.C.; resources, Y.C.; data curation, X.H. and X.Y.; writing—original draft preparation, Z.W.; writing—review and editing, Z.W., X.H. and D.Y.; visualization, Y.C.; supervision, D.Y. and X.Y.; project administration, D.Y. and X.Y.; funding acquisition, D.Y. and K.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by: 1. the National Natural Science Foundation of China (No. 62002227) and 2. the Zhejiang Provincial Natural Science Foundation of China (Grant No. LY20F020011).

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Not applicable.

Conflicts of Interest: The authors report no conflict of interest. All authors are responsible for the contents in and the writing of this article.

References

1. O'Connell, J.F.; Warnock-Smith, D. An investigation into traveler preferences and acceptance levels of airline ancillary revenues. *J. Air Transp. Manag.* **2013**, *33*, 12–21. [CrossRef]
2. Warnock-Smith, D.; O'Connell, J.F.; Maleki, M. An analysis of ongoing trends in airline ancillary revenues. *J. Air Transp. Manag.* **2017**, *64*, 42–54. [CrossRef]
3. Rouncivell, A.; Timmis, A.J.; Ison, S.G. Willingness to pay for preferred seat selection on UK domestic flights. *J. Air Transp. Manag.* **2018**, *70*, 57–61. [CrossRef]
4. Klislinar, E.; Widjaja, A.W. Analysis of Willingness to Pay for Ancillary Revenue of Full Service Airline (The Case of Garuda Indonesia). *KnE Soc. Sci.* **2020**, 1213–1230. [CrossRef]
5. Chiambaretto, P. Air passengers' willingness to pay for ancillary services on long-haul flights. *Transp. Res. Part E Logist. Transp. Rev.* **2021**, *147*, 102234. [CrossRef]
6. Zhao, G.; Cui, Y.; Cheng, S. Dynamic pricing of ancillary services based on passenger choice behavior. *J. Air Transp. Manag.* **2021**, *94*, 102058. [CrossRef]
7. Kolbeinsson, A.; Shukla, N.; Gupta, A.; Marla, L.; Yellepeddi, K. Galactic Air Improves Airline Ancillary Revenues with Dynamic Personalized Pricing. Available online: https://ssrn.com/abstract=3836941 (accessed on 13 January 2022). [CrossRef]
8. Suki, N.M. Passenger satisfaction with airline service quality in Malaysia: A structural equation modeling approach. *Res. Transp. Bus. Manag.* **2014**, *10*, 26–32. [CrossRef]
9. Scotti, D.; Dresner, M.; Martini, G. Baggage fees, operational performance and customer satisfaction in the US air transport industry. *J. Air Transp. Manag.* **2016**, *55*, 139–146. [CrossRef]
10. Nhamo, G.; Dube, K.; Chikodzi, D. *Counting the Cost of COVID-19 on the Global Tourism Industry*; Springer: Berlin/Heidelberg, Germany, 2020.
11. Maneenop, S.; Kotcharin, S. The impacts of COVID-19 on the global airline industry: An event study approach. *J. Air Transp. Manag.* **2020**, *89*, 101920. [CrossRef]
12. Vinod, B. Airline revenue planning and the COVID-19 pandemic. *J. Tour. Futur.* **2021**. doi: 10.1108/JTF-02-2021-0055. [CrossRef]
13. Xue, D.; Liu, Z.; Wang, B.; Yang, J. Impacts of COVID-19 on aircraft usage and fuel consumption: A case study on four Chinese international airports. *J. Air Transp. Manag.* **2021**, *95*, 102106. [CrossRef] [PubMed]
14. Gunardi, G.; Martono, M.R.S.H. COVID-19: The Impact on Air Transportation Tariff in Indonesia. In *International Conference on Economics, Business, Social, and Humanities (ICEBSH 2021)*; Atlantis Press: Paris, France, 2021; pp. 344–349.
15. Arena, M.; Aprea, C. Impact of Covid-19 Pandemic on Air Transport: Overview and Implications. *Adv. Env. Eng. Res.* **2021**, *2*, 2101002. [CrossRef]
16. Truong, D. Estimating the impact of COVID-19 on air travel in the medium and long term using neural network and Monte Carlo simulation. *J. Air Transp. Manag.* **2021**, *96*, 102126. [CrossRef]
17. Dube, K.; Nhamo, G.; Chikodzi, D. COVID-19 pandemic and prospects for recovery of the global aviation industry. *J. Air Transp. Manag.* **2021**, *92*, 102022. [CrossRef]
18. IdeaWorksCompany. The 2019 CarTrawler Yearbook of Ancillary Revenue. Available online: https://www.cartrawler.com/ct/ancillary-revenue/2019-cartrawler-ancillary-yearbook// (accessed on 17 September 2019).
19. IdeaWorksCompany. Ancillary Revenue Defined. Available online: https://ideaworkscompany.com/industry-definitions// (accessed on 26 December 2021).
20. Wittmer, A.; Rowley, E. Customer value of purchasable supplementary services: The case of a European full network carrier's economy class. *J. Air Transp. Manag.* **2014**, *34*, 17–23. [CrossRef]
21. Jeon, M.S.; Lee, J.H. Estimation of willingness-to-pay for premium economy class by type of service. *J. Air Transp. Manag.* **2020**, *84*, 101788. [CrossRef]
22. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 4768–4777.
23. Chen, C.F.; Wu, T.F. Exploring passenger preferences in airline service attributes: A note. *J. Air Transp. Manag.* **2009**, *15*, 52–53. [CrossRef]
24. Correia, A.; Pimpão, A.; Tão, M. Willingness to pay for frills when travelling with low-cost airlines. *Tour. Econ.* **2012**, *18*, 1161–1174. [CrossRef]
25. Han, H.; Lee, K.S.; Chua, B.L.; Lee, S.; Kim, W. Role of airline food quality, price reasonableness, image, satisfaction, and attachment in building re-flying intention. *Int. J. Hosp. Manag.* **2019**, *80*, 91–100. [CrossRef]
26. Shukla, N.; Kolbeinsson, A.; Marla, L.; Yellepeddi, K. From Average Customer to Individual Traveler: A Field Experiment in Airline Ancillary Pricing. Available online: https://ssrn.com/abstract=3518854 (accessed on 13 January 2022). [CrossRef]
27. Shaw, M.; Tiernan, S.; O'Connell, J.F.; Warnock-Smith, D.; Efthymiou, M. Third party ancillary revenues in the airline sector: An exploratory study. *J. Air Transp. Manag.* **2021**, *90*, 101936. [CrossRef]
28. Shao, S.; Kauermann, G.; Smith, M.S. Whether, when and which: Modelling advanced seat reservations by airline passengers. *Transp. Res. Part A Policy Pract.* **2020**, *132*, 490–514. [CrossRef]
29. Zhou, Y.; Zhang, T.; Mo, Y.; Huang, G. Willingness to pay for economy class seat selection: From a Chinese air consumer perspective. *Res. Transp. Bus. Manag.* **2020**, *37*, 100486. [CrossRef]

30. Yoon, M.; Lee, H. Seat assignment problem with the payable up-grade as an ancillary service of airlines. *Ann. Oper. Res.* **2021**, *307*, 1–15. [CrossRef]
31. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.Y. Lightgbm: A highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 3146–3154.
32. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
33. Benesty, J.; Chen, J.; Huang, Y.; Cohen, I. Pearson correlation coefficient. In *Noise Reduction in Speech Processing*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 1–4.
34. McHugh, M.L. The chi-square test of independence. *Biochem. Med.* **2013**, *23*, 143–149. [CrossRef] [PubMed]
35. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
36. Donoho, D.L. High-Dimensional Data Analysis: The Curses and Blessings of Dimensionality. *AMS Math Challenges Lect.* **2000**, *1*, 1–33.
37. Longadge, R.; Dongre, S. Class imbalance problem in data mining review. *arXiv* **2013**, arXiv:1305.1707.
38. Sarmanova, A.; Albayrak, S. Alleviating class imbalance problem in data mining. In Proceedings of the 2013 21st Signal Processing and Communications Applications Conference (SIU), Haspolat, Turkey, 24–26 April 2013; pp. 1–4.
39. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd Acm Sigkdd International Conference On Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
40. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]
41. Batista, G.E.; Bazzan, A.L.; Monard, M.C. Balancing Training Data for Automated Annotation of Keywords: A Case Study. Brazilian Workshop on Bioinformatics, Macaé, Brazil, 3–5 December 2003; pp. 10–18.
42. Shapley, L.S. *17. A Value for N-Person Games*; Princeton University Press: Princeton, NJ, USA, 2016.