Upgrade Optimization in the Airline Industry_A Privacy Preserving Federated Learning Approach

Upgrade Optimization in the Airline Industry:

A Privacy-Preserving Federated Learning Approach

Abstract

A key issue of making upgrade decisions is to match the most relevant upgrade offers to the right customers at the right time. To optimize upgrade strategies and profitability, companies seek to break “data silos” between themselves and other business partners for a more holistic view of customers’ consumption experiences. However, multi-source data fusion may lead to potential privacy leakage. To overcome these two challenges in data silos and privacy protection, this study introduced a privacy-preserving federated learning (FL) approach and explained the process of using FL in modeling airline passengers’ willingness to pay for upgrade offers. Using a case study of an airline company, this study demonstrated how FL-based upgrade algorithms using multi-source data can be developed to improve upgrade prediction accuracy while preserving customers’ personal data privacy. This study offers significant theoretical and practical implications for upgrade optimization in the contexts of airlines and other hospitality-related businesses.

Statement of Key Contributions

Upgrade optimization is a complicated data-driven process that requires companies to break data silos and utilize as much customer information as possible for generating business insights. However, given the potential risk of data leakage and customers’ privacy rights, companies must take measures to ensure personal data privacy while utilizing and analyzing the “big data” of customers. To overcome this dilemma of privacy protection and data silos, this study introduced the privacy-preserving federated learning approach as a promising solution for airline upgrade optimization. The proposed FL approach allows companies to store customer data locally and applies encryption methods to ensure personal data privacy during the entire process. This study contributes to the literature of upgrade optimization by introducing the new FL approach for developing machining learning models to predict customers' reaction to upgrade offers. Although we focus on the airline industry in our case study, the proposed FL approach can be applied to other industries with a similar issue of upgrade optimization such as hotels or cruise lines, and car rental.

From the practical perspective, the proposed FL approach provides airlines a novel solution for maximizing revenue and protecting customer privacy at the same time. Although the airline industry is among the early adopters of revenue management systems (Kimes, 2003), integrating large-scale multi-source data to optimize revenue management decisions is still a new domain for many airline companies. Especially, after the introduction of the General Data Protection Regulation (GDPR) and other similar privacy protection regulations, it becomes more urgent for airlines to consider the issue of privacy protection in their business analytics practice. Therefore, this study offers a timely and useful FL approach for airlines to work with other business partners to utilize customer data for drawing business insights cooperatively while ensuring personal data privacy.

Research Questions

Offering customers product or service upgrades is an important revenue management strategy that has been widely adopted in many businesses. A key issue of making upgrade decisions is to match the most relevant upgrade offers to the right customers at the right time for the most optimal price (Steinhardt & Gönsch, 2012). Not every customer is a good fit for an upgrade, and it is useless to push additional products or services on someone who doesn't truly need them (Yılmaz, Pekgün, & Ferguson, 2017). Therefore, identifying the customers who are likely to pay for the upgrade offers is critical in the upgrade decision-making process. In the airline industry, big data and business intelligence systems have been playing an important role in making revenue management decisions (Knorr, 2019; Krämer, Friesen,

& Shelton, 2018; Wittman & Belobaba, 2017). Traditionally, airline companies rely on the business intelligence systems to accumulate the passengers' travel history data, then use the collected internal data to estimate passengers' willingness to pay for an upgrade (Vinod, 2016). These internal data of airline companies, however, do not track passengers' overall tourism consumption with other airlines, hotels, or other tourism service providers. In other words, relying on airlines' internal data cannot reflect the bigger picture of passengers' overall buying power and the individual's willingness of paying for an upgrade (Vinod, 2016).

Thus, multi-source data fusion has become more and more popular in recent applications of business analytics (Meng & Du, 2016; Mangortey, Gilleron, Dard, Pinon-Fischer & Mavris, 2019). On the one hand, data fusion is an important process of integrating multiple data sources to produce more consistent and accurate information regarding customers' consumption profiles (Zheng, 2015). Since no single company owns all data about customers' daily consumption activities, multi-source data fusion enables companies to break the "data silos" between themselves and other business partners, which further provides a more holistic view for developing personalized products and precision marketing strategies (Lau et al., 2019; Zheng, 2015). On the other hand, since customers' data collected from multi-sources may include the sensitive private information, the data fusion process could also lead to potential privacy leakage (Gwebu & Barrows, 2020; Morris, Kleist, Dull, & Tanner, 2014). In recent years, the data security and privacy issues in big data analytics have attracted increasing attentions from academia, industrial practitioners, and governments (Chen & Fiscus, 2018; Hall & Ram, 2020; Line et al., 2020). For example, the European Union (EU) adopted the General Data Protection Regulation (GDPR) to protect personal data relating to individuals in the EU, which require businesses to comply with the GDPR rules when collecting personal data from consumers. To comply with these privacy protection regulations, companies must be very careful in collecting, sharing, and analyzing customers' personal data.

Being aware of the aforementioned issues in data silos and privacy protection, this study introduced a privacy-preserving federated learning (FL) approach for modeling airline passengers' willingness to pay for upgrade offers. Federated learning is a new confidential computing technique that allows companies to train a model cooperatively by exchanging model parameters instead of the actual raw data, which might include customers' privacy sensitive information (Li, Fan, Tse, & Lin, 2020). To protect personal data privacy, entities participating in a FL process can store their raw data locally and avoid sharing customer data with other entities in the FL networks. Entities of the FL networks regularly share their knowledge and parameters of training machine learning models for constantly improve the models' performance and accuracy (Li, Sahu, Talwalkar, & Smith, 2020). Using a case study of a Chinese airline company, this study demonstrated how FL-based upgrade models using multi-source data can be developed to improve the accuracy of predicting customers' willingness to pay for upgrades while preserving customers' personal data privacy. Although we focus on the airline industry in our case study, the proposed FL approach can be applied to other industries with a similar issue of service upgrade optimization such as hotels or cruise lines. This study offers significant theoretical and practical implications for upgrade optimization in the contexts of airlines and other hospitality-related businesses.

Case Study of Airline Upgrade Optimization

To illustrate the proposed PL approach, a case study of a Chinese airline company was conducted to develop machine learning models for predicting passengers' willingness to pay for upgrade offers. To protect the company's commercial confidentiality, it is named "Airlines A" in this study. In China, the civil aviation market has been opened up to private investors since 2005. During the past decades, more and more new private airlines joined the market, so that the competition among airline companies has become increasingly intensive. Therefore, upgrade optimization is an important revenue management strategy for the success of airline companies in such a competitive market. In the following case study, we adopt the vertical federated learning approach for modeling passengers' willingness to pay for upgrades. The following section describes the process of proposed vertical federated learning approach.

Data preparation

The data sources used in this case study include the internal database of Airlines A and the external dataset from another big data service provider (Data Company B) in China. The Airlines A's database includes 19 variables of travel-related information for 13,546 passengers, such as passengers' ID card number, number of domestic flights in the last year, number of flight delay insurance purchases, and loyalty program membership level. The database of Data Company B consists of 26 variables relevant to individuals' social economic background and travel experience, such as annual income range, ownership of cars, parenthood status, and marriage status. This external database includes records for 47,159 passengers. All data were collected in the past three years. The variables of each data sources were listed below.

(1) Variables form Airlines A Database:

- *Dependent Variables:* Responses to upgrade offers (pay for the upgrade=1; reject to pay for the upgrade=0)
- *Independent Variables:* Passengers' ID card number; mobile number; Gender; Number of domestic flights in the last year; Proportion of domestic economy class in the last year; Proportion of domestic business class or first class in the last year; Average discount rate in recent years; Number of economy-class flights in the last year; Economy class proportion in the last year; Number of domestic business or first-class flights in the last year; Number of National Day flights in recent three

years; Number of May 1 flights in recent three years; Number of Spring Festival flights in recent three years; Flights in recent three years; Flights in recent two years; Number of business or first-class domestic flights in the last year; Number of complaints in recent year; Number of advance seat selection in the last year; Number of flight delay insurance purchases in recent two years; loyalty program membership level.

(2) Variables from Big Data Company B:

- *Independent Variables:* ID card number; telephone number; Annual income range; ownership of cars; Parenthood status; Marriage status; Education levels; Ownership of residential property; Hotel brand preference; peer-to-peer accommodation preference; Preference of tourism scenic spots; Consumption preference for luxury brands; Consumption preference for mass market brands.

This case study focuses on predicting passengers' willingness to pay for upgrade offers, we set responses to upgrades as the dependent variable (pay for the upgrade=1; reject to pay for the upgrade=0). An acceptance of an upgrade offer was marked as 1, while a rejection was mark as 0. To protect the passengers' sensitive information, the anonymization technique was applied during the data preparation process.

Data Analysis

To apply the vertical FL approach, we adopted a federated intelligence system known as the FL.Insight platform (http://www.techvalley.com.cn/tech/50), which serves as a data aggregator to merges the model parameters (not the raw data) of Airlines A and Big Data Company B. Figure 1 depicts the data analysis process of the FL.Insight platform.

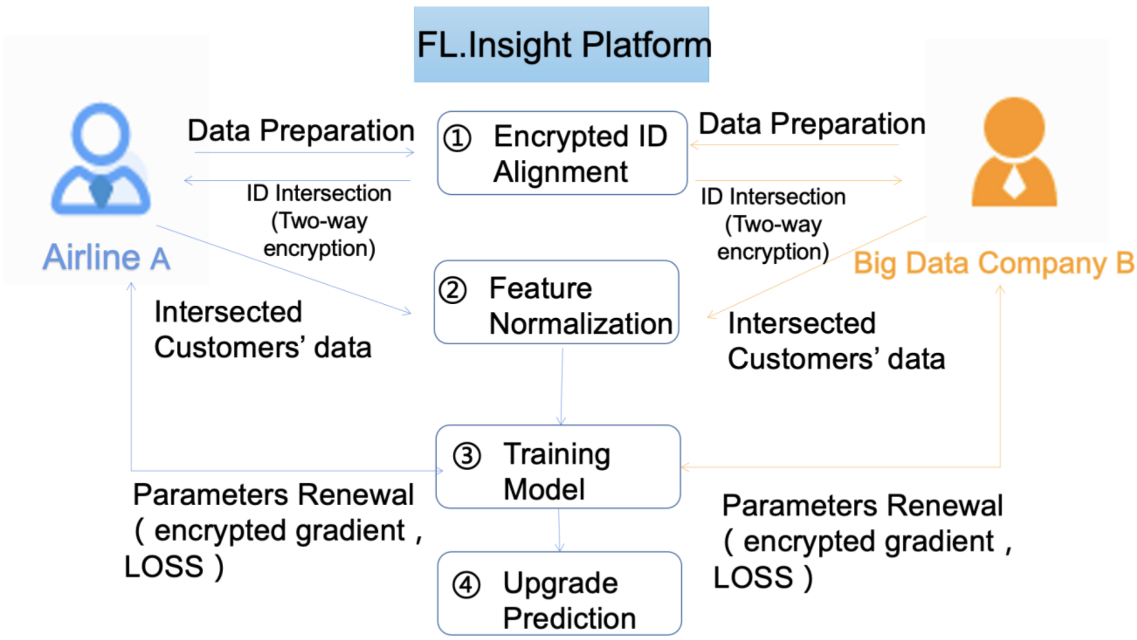

Figure 1: The Flow Chat of FL. Insight Platform for Airline Upgrade Optimization

Given the decision of paying for upgrade offers is a binary classification issue, there are three machine learning algorithms suitable for solving binary classification problems: logistic regression, SecureBoost, and neural network. Each machine learning algorithms have its pros and cons. First, logistic regression has good business explanation, consumes less resources, and can get the so-called probability prediction value. The speed of the updating model is fast. However, when the feature space is very large, the performance of logical regression is not very good, which depends on the conditions of hardware resources. Second, SecureBoost achieves the minimum classifying loss by finding a set of best decision trees having the relatively low model complexity. Generally speaking, the features of nodes are pre-sorted before iterations, and the optimal segmentation points are selected by traversing. When the amount of data is large, the algorithm takes a long time, and the interpretability of the model is relatively poor. Third, the performance of the neural network model is better than almost other machine learning algorithms when the amount of data and dimensions are large enough. But it also has some disadvantages. For example, a large amount of data support is needed. The high computational cost and poor interpretability of the model are also problems.

Given the three algorithms have their own advantages, it is hard to tell at the beginning which algorithm is the most suitable for predicting passengers' upgrade willingness. Therefore, we adopted all three algorithms for modeling passengers' upgrade willingness, then compared the prediction performance of each developed model to identify the most suitable one. For the different purposes of model development and comparison, we divided the datasets into training subset and testing subset. The training subset includes records of 9,472 passengers with matching IDs in the databases of Airlines A and Data Company B, while the testing subset incudes records of 2,753 passengers with matching IDs in both databases.

To evaluate the advantage of privacy-preserving multi-source data fusion, we also conducted unilateral logistic regression analysis using the single source data from the Airlines X only, then compared the model's performance with the other three models using multi-source data.

Last, to evaluate the performance of the developed models, the recall rate, accuracy, precision, F0.5-score and the area under the curve (AUC) indexes were used.

(1) Upgrade recall rate:

$$\mathrm{Re}call = \frac{TP}{TP + FN}$$

TP: The number of customers who were predicted to upgrade and actually upgraded.

FN: The number of customers predicted not to upgrade but actually upgraded.

The proportion of upgraded customers that are correctly predicted, i.e., the number of customers with predicted upgrades and real upgrades/the total number of customers with real upgrades. The higher the upgrade recall rate is, the model can predict the real upgrade customers more accurately and therefore improve the revenue.

(2) Upgrade accuracy index:

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

TP: Number of customers who were forecasted upgrades and actually upgraded.

TN: The number of customers who were not expected to upgrade and did not upgrade.

FP: Number of customers who were predicted to upgrade but did not upgrade.

FN: Number of customers who actually upgraded without being expected to upgrade.

The ratio of real upgrade customers and real non-upgrade customers that are correctly predicted, i.e., the correct prediction of upgrades/total customers. The higher the accuracy of upgrade prediction, the better the model can judge whether the customer has demand for upgrade, and more accurate upgrading marketing.

(3) Upgrade precision index:

$$\text{Precision} = \frac{TP}{TP + FP}$$

TP: Number of customers who were forecasted upgrade and actually upgrade.

FP: Number of customers who were predicted to upgrade but did not upgrade.

Proportion of customers who are forecasted upgrade and actually upgraded to total customers who forecast upgrades, i.e., the number of customers who were forecasted to upgrades and actually upgraded/the total number of customers who were forecasted upgrade. The higher the accuracy rate of upgrade prediction, the better the model can predict upgrading demand, and then carry out more accurate upgrading marketing.

(4) F0.5-Score:

$$\text{F-Score} = \frac{(1 + 0.5) * P * R}{0.5 * P + R}$$

P: Upgrade Precision.

R: Upgrade Recall.

The F-score is a measure of the classification problem and the F0.5-score considers Recall to be half as important as precision. Therefore, F0.5-Score is an indicator to comprehensively measure the upgrade recall rate and upgrade precision rate.

(5) Area under the curve (AUC) index: For a randomly given group of upgraded and non-upgraded customers, the probability that the model's predicted value of upgraded customers is greater than that of non-upgraded customers. The larger the AUC, the higher the model's degree of distinction between positive and negative samples, and hence, the better the model's classification performance.

Summary of Findings

Firstly, unilateral modeling is conducted on the data of the airline company to facilitate comparison with the effect of federated modeling. The results show that the upgrade recall rate of the unilateral model which uses the test set of airline data is 124/200=0.62. The upgrade accuracy is (124+1827)/(200+2553)=0.709. The upgrade precision is 124/(124+726)=0.146. The F-score is ((1+0.5）*0.62*0.146）/(0.5*0.146+0.62)=0.196. AUC index is 0.746.

Next, we join the big data company dataset for federated modeling. The results show that the upgrade recall increases to 189 / 200 = 0.945, which is 52 % higher than that of the local unilateral modeling. The upgrade accuracy also increases to (189+2522) / (200+2553) =0.985, which has a 39% improvement. The upgrade precision is 189 / (189+31)=0.850; F-score is ((1+0.5) * 0.945 * 0.85) / (0.5 * 0.0.85+0.945) =0.879 , and the AUC increase to 0.979, 31% improvement in model performance over 0.735 of local one-sided modeling.

Then, the federated SecureBoost modeling was conducted. After adding the features of big data companies to the joint model, the recall rate of the test set increases to 190 / 200 = 0.95, which is 53 % higher than that of the local unilateral modeling. The accuracy of cabin upgrading increases to (190 + 2543) / (200 + 2553) = 0.993, which increases by 40 %. The precision ratio of upgrade is 190 / (190 + 10) = 0.95. F-score is (1 + 0.5) * 0.95 * 0.9) / (0.5 * 0.95 + 0.95) = 0.950. The AUC also increases to 0.972, which is 30 % higher than that of 0.735 in local unilateral modeling.

In addition, we evaluated the results of the neural network model. After adding the features of the big data company to the joint model, the recall rate of the test set is 66 / 200 = 0.33, the accuracy rate of the upgrade is (66 + 656) / (200 + 2553) = 0.262, the precision rate of the upgrade is 66 / (66 + 1897) = 0.03, the F-score is (1 + 0.5) * 0.33 * 0.03) / (0.5 * 0.03 + 0.33) = 0.043, and the AUC index is 0.254. The performance is poor.

To evaluate the performance of each model, Table 1 summarized the key indexes of all models in this study. As shown in Table 1, the SecureBoost model has high values of above 0.9 for all key indexes, indicating the best overall performance than other models. The key indexes of federated logistic regression model are also above 0.85, which ranks the second after the SecureBoost model. The performance of neural network model is relatively poorer because the neural network algorithm requires a large-scale dataset to support its accuracy. Given the dataset size in this case study is not large enough, the performance of neural network model is poorer than other models. However, comparing with traditional unilateral model using single-source data, the federated logistic regression and SecureBoost models demonstrate better model performance. This indicates that the proposed FL approach can enhance the accuracy of modeling airline passengers' willingness to pay for upgrade offers while preserving passengers' data privacy.

Table 1 Model Performance Comparison

Model	Recall of upgrade	Accuracy of upgrade	Precision of upgrade	AUC	F-score	Running time
Unilateral Modeling	0.620	0.709	0.146	0.746	0.196	19m
Federated SecureBoost	0.950	0.993	0.950	0.972	0.950	36m
Federated	0.945	0.985	0.850	0.979	0.879	1h1m

Logistic Regression						
Federated Neural Network	0.330	0.262	0.030	0.254	0.043	1h27m

It's worthy to note that the FL-based models generally took longer running time than the traditional unilateral model due to the design of FL approach in ensuring data privacy. Since all raw data of all participants cannot not be disclosed in the FL process, the intermediate results of the data information cannot be extrapolated. So, each iteration of the model requires strict decrypted training, as well as the encryption and upload steps. At the same time, because the data cannot be gathered together, participants need to interact with arbiter in each iteration. Also, as the 1024-bit or 2048-bit Paillier homomorphic encryption algorithm is used in federated learning encryption, the transmitted data increase from a few bytes to thousands of bytes, which directly leads to a significant decrease in the training speed of the federated learning system comparing to the traditional machine learning. At the same time, due to the large number of model data that need to be exchanged with each other in the training process for the algorithms, the stability of cross-network data transmission needs to be taken into account in real-world practice, which also leads to slower training speed of the FL approach. For companies' business operation, however, the risk of data leakage and the legal concerns of complying to privacy protection laws are much more important issues than the increased running time of data processing. Overall, the FL approach is a good solution for companies to ensure both modeling accuracy and privacy protection.

Selected References

Chen, Hsiangting Shatina, and Joseph Fiscus (2018). "The inhospitable vulnerability: A need for cybersecurity risk assessment in the hospitality industry." *Journal of Hospitality and Tourism Technology*, 9(2), 223-234.

Gwebu, Kholekile, and Clayton W. Barrows. (2020). “Data breaches in hospitality: is the industry different?.” *Journal of Hospitality and Tourism Technology,* 11(3), 511-527.

Hall, Colin Michael, and Yael Ram (2020). “Protecting privacy in tourism–a perspective article.” *Tourism Review*, 75(1), 76-80.

Knorr, Andreas (2019). “Big Data, Customer Relationship and Revenue Management in the Airline Industry: What Future Role for Frequent Flyer Programs?.” *Review of Integrative Business and Economics Research*, 8(2), 38-51.

Krämer, Andreas, Mark Friesen, and Tom Shelton (2018). “Are airline passengers ready for personalized dynamic pricing? A study of German consumers.” *Journal of Revenue and Pricing Management,* 17(2), 115-120.

Lau, Billy Pik Lik, Sumudu Hasala Marakkalage, Yuren Zhou, Naveed Ul Hassan, Chau Yuen, Meng Zhang, and U-Xuan Tan (2019). “A survey of data fusion in smart city applications.” *Information Fusion*, 52, 357-374.

Li, Li, Yuxi Fan, Mike Tse, and Kuo-Yi Lin (2020). “A review of applications in federated learning.” *Computers & Industrial Engineering,* 106854.

Li, Tian, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith (2020). “Federated learning: Challenges, methods, and future directions.” *IEEE Signal Processing Magazine*, 37(3), 50-60.

Line, Nathaniel D., Tarik Dogru, Dahlia El-Manstrly, Alex Buoye, Ed Malthouse, and Jay Kandampully (2020). “Control, use and ownership of big data: A reciprocal view of customer big data value in the hospitality and tourism industry.” *Tourism Management*, 80, 104106.

Mangortey, Eugene, Jerome Gilleron, Ghislain Dard, Olivia J. Pinon-Fischer, and Dimitri N. Mavris (2019). “Development of a data fusion framework to support the analysis of aviation big data.” Paper presented at AIAA Scitech 2019 Forum, San Diego, California (January 7-11).

Xiaofeng, Meng, and Du Zhijuan (2016). “Research on the big data fusion: issues and challenges.” *Journal of Computer Research and Development*, 53(2), 231.

Morris, Bonnie W., Virginia Franke Kleist, Richard B. Dull, and Cynthia D. Tanner (2014). “Secure information market: A model to support information sharing, data fusion, privacy, and decisions.” *Journal of Information Systems*, 28(1), 269-285.

Steinhardt, Claudius, and Jochen Gönsch (2012). “Integrated revenue management approaches for capacity control with planned upgrades.” *European Journal of Operational Research*, 223(2), 380-391.

Vinod, Ben (2016). “Evolution of yield management in travel.” *Journal of Revenue and Pricing Management,* 15(3), 203-211.

Wittman, Michael D., and Peter P. Belobaba (2017). “Personalization in airline revenue management–Heuristics for real-time adjustment of availability and fares.” *Journal of Revenue and Pricing Management,* 16(4), 376-396.

Yılmaz, Övünç, Pelin Pekgün, and Mark Ferguson (2017). “Would you like to upgrade to a premium room? Evaluating the benefit of offering standby upgrades.” *Manufacturing & Service Operations Management,* 19(1), 1-18.

Zheng, Yu (2015). “Methodologies for cross-domain data fusion: An overview.” *IEEE transactions on big data,* 1(1), 16-34.