



# A two stage machine learning approach for Modeling Customer Lifetime Value in the Chinese Airline Industry

## Document Version

Final published version

[Link to publication record in Manchester Research Explorer](#)

## Citation for published version (APA):

Chen, S., Yinghua Huang, Xu, D-L., & Jiang Wei (2020). A two stage machine learning approach for Modeling Customer Lifetime Value in the Chinese Airline Industry. In S. Blanchard, A. Epp, & G. Mallapragada (Eds.), *AMA Summer Academic Conference 2020 : Bridging Gaps Marketing in an Age of Disruption PROCEEDINGS* (Vol. 31, pp. 1018-1021). American Marketing Association (AMA).

## Published in:

AMA Summer Academic Conference 2020

## Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

## General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

## Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact [uml.scholarlycommunications@manchester.ac.uk](mailto:uml.scholarlycommunications@manchester.ac.uk) providing relevant details, so we can investigate your claim.



**A Two-Stage Machine Learning Approach for Modeling Customer Lifetime Value  
in the Chinese Airline Industry**

Sien Chen<sup>1,2,3</sup>

1. Alliance Manchester Business School,  
University of Manchester, Manchester, UK
2. Research Institute of Internet Industry,  
Tsinghua University, Beijing, China
3. Antai College of Economics & Management, Shanghai Jiao Tong University, Shanghai, China  
[sien.chen@postgrad.manchester.ac.uk](mailto:sien.chen@postgrad.manchester.ac.uk)

Yinghua Huang\*

Dept. of Hospitality, Tourism and Event Management  
San Jose State University  
San Jose, USA  
[yinghua.huang@sjsu.edu](mailto:yinghua.huang@sjsu.edu)

Dong-Ling Xu

Alliance Manchester Business School  
University of Manchester  
Manchester, UK  
[Ling.Xu@manchester.ac.uk](mailto:Ling.Xu@manchester.ac.uk)

Jiang Wei

Antai College of Economics & Management  
Shanghai Jiao Tong University  
Shanghai, China  
[jiangwei@sjtu.edu.cn](mailto:jiangwei@sjtu.edu.cn)

\*corresponding author

# **A Two-Stage Machine Learning Approach for Modeling Customer Lifetime Value in the Chinese Airline Industry**

## **Abstract**

This study proposed a two-stage machine learning approach for modeling customer lifetime value (CLV) by integrating the advantages of three traditional machine learning methods (Stage 1) and an innovative method of evidence reasoning (ER) modeling (Stage 2). Unlike previous studies of developing a single CLV model, this study adopted the ER approach (Stage 2) to fuse the results obtained from the Stage 1 (i.e., logistic regression, gradient boosting decision tree, and neural network models) to improve the prediction accuracy. We illustrated the proposed two-stage approach using a case study of Chinese airline passengers. We analyzed a dataset of over 100,000 Chinese airline passengers, which includes 327 variables related to passengers' travel experience with major airline companies and online travel agents in China. The case study showed that the proposed two-stage approach performed better in comparison with other machine learning algorithms. This study provided airline companies a useful data-driven method to identify high-value passengers and allowed airlines to allocate the CRM resources more effectively.

**Keywords:** machine learning; evidential reasoning; customer lifetime value; airline industry

## **Statement of Key Contributions**

This study is the first, to the best of our knowledge, to integrate multiple machine learning methods and network-wide data sources for modeling CLV in the airline industry. Unlike previous studies of developing a single CLV model, this study applied the evidence reasoning method (Stage 2) to integrate the results of multiple machine learning models (Stage 1). The evidence reasoning method has an advantage of fusing data and generating classification decisions when facing uncertainty or

conflicts in multiple pieces of information (Yang et al., 2018). Therefore, the proposed two-stage approach can overcome the limitation of relying on a single model and further enhance the accuracy of CLV assessment.

The existing CLV models primarily focus on customers' consumption history with a subject company. However, these models fail to examine how customers buy products or services from the competitors of the subject company (Castéran, Meyer-Waarden, & Reinartz, 2017). In this study, we used a large-scale network-wide dataset that combined an airline company's internal data and a third-party's network-wide data. The network-wide data source integrated passengers' travel experience with major airline companies and online travel agents in China. Therefore, this study modeled passengers' CLV based on their consumption across the entire network of airline ticket distribution. This study contributes to the literature by overcoming the limitation of modeling CLV based on customers' experience with a single company, so that we could see the bigger picture of passengers' overall consumption patterns and identify high-value customers more accurately.

As demonstrated in the case study, the proposed two-stage approach can help airline companies to segment high-value passengers more precisely, so that airlines can develop personalized CRM strategies to improve and maintain customer relationships more wisely. Given the current competition among airline companies are very intensive, this study offers very important managerial implications for airlines.

# **A Two-Stage Machine Learning Approach for Modeling Customer Lifetime Value in the Chinese Airline Industry**

## **1. Introduction**

Customer lifetime value (CLV) is an important concept in the field of customer relationship management (CRM) because CLV indicates a customer's profitability over the person's life of relationship with a company (Dwyer, 1989). Using CLV as a key metric, CRM managers can evaluate current, and potential customers' profitability, then customize products and services to acquire and retain profitable customers (Stone, Woodcock, & Wilson, 1996).

Given the importance and popularity of CLV, marketing researchers have developed various models to estimate CLV over past decades, including the well-known Recency-Frequency-Monetary (RFM) model, the Pareto/NBD model and so on. These existing models primarily assess CLV based on customers' consumption history with a subject company. However, they fail to examine how customers buy products or services from the competitors of the subject company (Castéran, Meyer-Waarden, & Reinartz, 2017). Therefore, the current literature of CLV modeling has a problem that we cannot see the bigger picture of customers' overall consumption experience with a company and its competitors due to the issues of data silo and unknown share of wallet, which leads to mistakes in identifying high-value customers.

Many researchers and industry practitioners, therefore, are calling for new modeling approaches to improve the accuracy of CLV assessment (Gupta et al., 2006). Previous studies have shown that the CLV models vary depending on the industry and characteristics of products and services (Pfeifer, Haskins, & Conroy, 2005; Venkatesan & Kumar, 2004). In this study, we focus on the airline industry and propose a two-stage machine learning approach for modeling the CLV of airline passengers. The current competition among airline companies is fierce, and the airlines are constantly facing the challenge of perishable inventory. In order to acquire and retain "right" passengers with a high CLV, many airline

companies have adopted the frequent flyer membership program to segment customers based on the passenger's travel history such as flight frequency, travel mileage, expenses and so on (Araujo & Kjellberg, 2015). Such traditional membership programs, however, only use the internal data from the airline company itself, which means that it cannot accurately measure the passenger's overall buying power and the person's potential of profitability in the future (Vinod, 2016).

To overcome the limitation discussed above, this study seeks to propose a two-stage machine learning approach to assess the airline passenger's CLV based on both the airline company's internal data and a third-party's network-wide data. First, we evaluated the information value (IV) of each variable related to passengers' travel experience and selected those features with a higher power of prediction for the next step. Then, we applied three machine learning models (i.e., logistic regression, gradient boosting decision tree, and neural network data mining) in Stage 1 and compared the results of each model in predicting passenger's CLV (high vs. low). In Stage 2, we used the evidential reasoning (ER) method to fuse the data obtained from Stage 1 to further improve the prediction accuracy. The ER method has an advantage of fusing data and generating classification decisions when facing uncertainty or conflicts in multiple pieces of information (Yang et al., 2018). This study is the first, to the best of our knowledge, to integrate multiple machine learning methods and network-wide data sources for modeling CLV in the airline industry.

To illustrate the proposed two-stage approach, we conducted a case study of Chinese airline passengers. We analyzed a dataset of over 100,000 Chinese airline passengers, which includes 327 variables related to passengers' travel experience with major airline companies and online travel agents in China. The results of the case study show that the proposed two-stage approach performs better in comparison with other machine learning algorithms. Therefore, this study provides the airline companies a useful data-driven approach to identify high-value passengers and allows the airlines to allocate the

CRM resources more effectively. Moreover, the proposed approach could be applied to assess CLV in other different industries and benefit the CRM practices of other companies.

## **2. Literature Review**

### *2.1. CLV Assessment*

CLV is defined as “the net profit or loss to the firm from a customer over the entire life of transactions of that customer with the firm” (Jain, Dipak, & Siddhartha, 2002). CLV assessment is an important topic in the CRM literature because it is the foundation of identifying high-value customers and segmenting customer groups. By measuring the CLV, the companies can distinguish between customers that are profitable, unprofitable, and have the potential to be profitable.

In the CRM literature, many models have been developed for measuring CLV, including Gupta’s CLV model (Gupta, Lehmann, & Stuart, 2003), FRM model (Hughes, 1996), NBD-Pareto model (Schmittlein, Morrison, & Colombo, 1987), Persistence model (Yoo & Hanssens, 2005), and Diffusion/Growth Models (Gupta, Lehmann, & Stuart, 2003). Although these existing models have been applied to assess CLV on various occasions, they have some limitations on their type of data used and the type of business the model is applicable to. One of the limitations is that the existing models only use internal data of the company to calculate the CLV and lack of external data for more comprehensive evaluation. Moreover, most CLV models focus on the cash flow stream from the customer but do not include variables related to customer’s consumption behavior (Jain, Dipak, & Siddhartha, 2002). Therefore, marketing researchers call for more studies to accurately predicting CLV based on a wide range of internal and external data, as well as variables related to the history of customer consumption.

### *2.2. Machine Learning Models*

Machine learning is a program that lets a computer "learn" people's learning process. The essence of machine learning is allowing the machine to simulate the process of human learning autonomously

based on the massive Internet data and the system powerful parallel computing power, and to make intelligent decision-making behavior by constantly “learning” the data. Machine learning is divided into supervised learning, unsupervised learning, and reinforcement learning (Jordan & Mitchell, 2015).

In supervised learning, each sample consists of an input object and the desired output (label). For example, "linear regression" in regression algorithms is to give a labeled data feature to try to learn a linear model to predict the real output value as accurately as possible (Schuld, Sinayskiy, & Petruccione, 2016). Linear regression and logistic regression algorithms are commonly used in regression algorithms for supervised learning. Classification in supervised learning refers to obtaining a classification function or classification model to map data objects of unknown categories into a given category (Armitage & Ober, 2010). Unlike regression, the predicted output value of the classification is a discrete category, and the regression is the continuous category. Currently, common classification algorithms are linear regression, logistic regression, Bayesian networks, decision tree, and so on (Kotsiantis, 2007). In this paper, we used logistic regression, gradient boosting decision tree (GBDT) and neural networks to conduct discrete analysis to estimate the CLV for airlines.

### *2.3. The ER method*

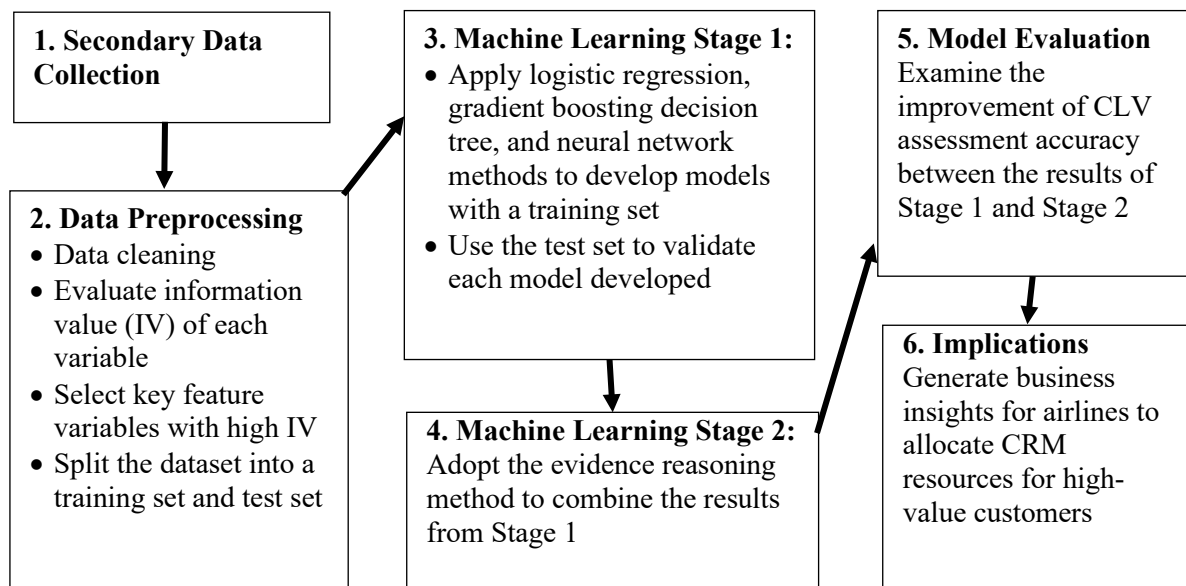
The ER method offers a framework to deal with classification imprecision issues from the perspective of uncertain information fusion (Yang, Xu, Yang, & Chen, 2018). Researchers introduced the ER approach to advance the Dempster rule and Bayes rule for evidence combination and information fusion (Xu & Yang, 2001). It has been incorporated into multi-attribute decision-making methods, forming the prototype of evidence reasoning which inherits the Dempster combination rule of evidence theory (Dempster, 1967) and gives the recursive algorithm of the overall synthesis (Yang & Singh, 1994). Due to the diversity of decision information, Yang and Xu (2001) introduced the information conversion method based on rules and utility by the original and established a general decision model based on



evidence reasoning, which can handle the decision problems of multiple data forms. The ER method has recently been applied to performance evaluation (Rahman, Mahmud, Pota, & Hossain, 2014), risk analysis (Tang & Nurmaya, 2011), inventory management (Atkinson et al., 2010), system prediction (Si, Hu, Yang, & Zhou, 2011), and work selection (Mahmud, Rahman, & Hossain, 2013). In this study, we applied the ER approach to investigate the assessment of CLV in the airline industry.

### 3. Research Framework

To assess airline passengers' CLV, this study proposed a two-stage machine learning approach by integrating the advantages of three traditional machine learning methods (Stage 1) and a new method of evidence reasoning modeling (Stage 2). Figure 1 depicts the framework of the proposed approach. The following sections describe the details of each step.



**Figure 1:** The proposed two-stage approach

#### 3.1. Data Collection

We collected data from two sources: a Chinese airline company and TravelSky. To protect the passengers' sensitive information, the anonymization technique was applied during the data collection process. The airline company provided 327 variables of passenger attributes such as membership status,

the number of bookings, the number of tickets, the loyalty program points, and so on. These attributes for 107,538 passengers were collected as independent variables.

We also collected the passenger's CLV score of the 107,538 passengers from TravelSky.

TravelSky is the largest travel service distribution network in China, which has partnered with 40 Chinese commercial airlines and over 350 foreign and regional commercial airlines. As such, TravelSky accumulated the network-wide data of passenger travel history and modeled passenger's CLV based on the passenger's overall travel experiences with multiple airline companies. The passengers' CLV score by TravelSky is a number ranging from 0 to 100. The higher the score, the higher the CLV of a passenger. Therefore, the TravelSky's CLV score is a good index to reflect a passenger's overall value, and we used it as the dependent variable.

### *3.2. Data Processing*

The independent variables from the airline company include both classification variables and continuous variables. Classification variables consist of both texture variables and numerical variables. The text variables were converted to an integer format. The original CLV score from TravelSky is a continuous variable ranging from 0 and 100. Then, we binarized the score into two categories: 0 (low-value passengers) and 1 (high-value passengers).

### *3.3. Feature Selection*

Feature selection is required because the number of initial features from the airline company is large, and some features may be irrelevant to the passenger's CLV (Yang, Xu, Yang, & Chen, 2018). Therefore, we used a feature selection procedure to remove irrelevant or redundant features and reduce the number of practical features to enhance model accuracy. In this study, we select features by using the information value (IV) indicator, which derives from the Information Theory to express the power of determination of each feature (Kullback, 1997). The IV indicator has been widely used to encode and

predict the input variables, which indicates the strength of the predictive ability of the variable (Dedu & Ganea, 2009; Yokota & Thompson, 2004). The formula for calculating IV is as follows:

$$IV = \sum(X_{dist} - \bar{X}_{dist}) \times \ln\left(\frac{X_{dist}}{\bar{X}_{dist}}\right) \quad (1)$$

$$X_{dist} = f(Y = y_k | X = x_i)$$

$$\bar{X}_{dist} = f(Y \neq y_k | X = x_i)$$

In this study,  $IV > 0.3$  was used as the criterion for feature selection.

### 3.4. Stage 1: Modeling CLV Using Three Machine Learning Algorithms

After the key features were selected, we divided the dataset into a training set and test set. The training set received records of 70% of the total passengers, while the test set received records of the other 30% passengers. Then, we adopted three machine learning algorithms (i.e., logistic regression, GBDT, and neural network) to conduct a binary classification for identifying passengers with high and low CLVs. The training set was used for constructing the three models. Then the test set was used for assessing the predictive power of these constructed models.

### 3.5. Stage 2: Model Fusion with ER

In order to fuse the results of the three machine learning models above, we adopted the ER method to improve prediction accuracy. The ER method is developed on the basis of evidence theory, overcoming the paradox of evidence theory (Wang, Yang, Xu, & Chin, 2006; Wang, Tang, & Wang, 2014). Due to the diversity of decision information, Yang et al. (2006) introduced the information conversion method based on rules and utility on the basis of the original and established a general decision model based on evidence reasoning, which can handle the decision problems of multiple data forms. The details of ER models can be found in Yang et al. (2006).

### 3.6. Model Evaluation

To evaluate the performance of each machine learning method, we measured the accuracy, precision, recall, and area under the curve (AUC) indicators for result comparison. For the ER method,

the mean square error (MSE) and the AUC indicator were used to measure the model performance. Then, we would identify a machine learning model with the best prediction performance.

### *3.7. Business Implications*

The results of machine learning models could generate business insights for airlines to identify high-value passengers more accurately. Then, the company can allocate CRM resources more wisely to acquire and maintain high-value customers.

## **4. A Case Study**

To illustrate the proposed approach, a case study of a Chinese airline company was conducted to predict its passengers' CLV. To protect the company's commercial confidentiality, it is named "Airlines X" in this study. Python 3.0 version was used for data analysis.

### *4.1. Data Collection*

We obtained travel experience data of 107,538 passengers from Airlines X. To protect passengers' sensitive information, we applied the data anonymization technique during the data collection process. This dataset consisted of 327 attributes related to four dimensions of the travel experience: passenger profile, travel history, route characteristics, and loyalty program status. The passenger profile variables included age, gender, booking channels, and so on. The travel history variables included the number of flights, number of ticket changes, number of reservation cancelations, flight schedule, and so on. The dimension of route characteristics includes travel destinations, departure cities, seat preference, number of delayed flights, and so on. The loyalty program variables included membership duration, point redemption time, points reward channels, points change over the past three months, and so on. All of these attributes were used as independent variables. In addition, we obtained the CLV score of these 107,538 passengers from TravelSky, which was used as the dependent variable.

#### 4.2. Data Processing

We used the binarization technique to transform the CLV score into 0 (low-value passengers) and 1 (high-value passengers). After examining the data distribution, we observed that passengers with a score between 99 and 100 accounted for around 20%, and the other 80% passengers had scores ranging from 0 to 98. Therefore, we selected 98 points as the boundary value, and the dependent variable binary is divided into two categories: 0-98 and 99-100. The scores ranging from 0 to 98 were transformed to 0, which indicated a relatively low CLV. The scores between 99 and 100 were recorded as 1, which indicated a high CLV.

#### 4.3. Feature selection

We calculated the IV for all dependent variables, then selected 60 attribute variables out of the 327 attributes using  $IV > 0.3$  as the criteria for feature selection.

#### 4.4. Stage 1: Modeling Results and Model Evaluation

In the first stage of machine learning, we adopted three machine learning algorithms (i.e., logistic regression, GBDT, and neural network) to conduct a binary classification for identifying passengers with high and low CLVs. We evaluated the effect of different thresholds on each model separately, setting the threshold from 0.1 to 0.9 with an increment of 0.1. After examining the degree of precision improvement by thresholds, we chose 0.8 as the optimal threshold for prediction. The reason was that the precision index was high when the threshold was 0.8, and the recall and accuracy were not too low as well. Table 1 summarized the metrics of accuracy, precision, recall, and AUC for each model.

Table 1: Model Performance Evaluation

Metrics	Machine Learning Stage 1			Stage 2
	logistic regression model	GBDT	neural network model	ER-based model fusion
Accuracy	0.775	0.818	0.810	0.834

Precision	0.826	0.840	0.845	0.827
Recall	0.774	0.847	0.823	0.904
AUC	0.863	0.893	0.888	0.897

---

#### 4.5. Stage 2: ER-based model fusion

To further improve the prediction accuracy, the ER method was used to fuse the results of Stage 1. To evaluate the performance of such model fusion, we compared the results of Stage 1 and Stage 2 in two ways.

First, the MSE of each model was calculated. In order to minimize MSE, we determined the weights of the three models in Stage 1. The weight of the logistic regression was 0, while the weights of GBDT and neural networks were 46% and 52% respectively. Then, using the ER rules, the results of GBDT were combined with the results of the neural network. Then, we obtained the minimum MSE for all models as shown in Figure 2. The MSE of the ER-based model fusion was 0.166, which was smaller than the MSEs of the other three models. Therefore, this result indicated that the ER-based model fusion performed better than the other three models.

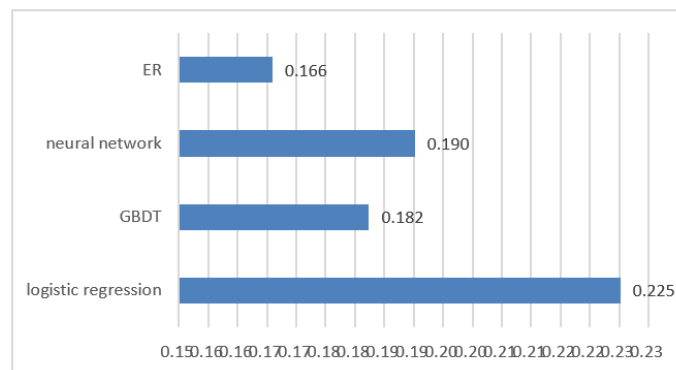


Figure 2: Comparison of Mean Squared Error

Second, we compared the assessment metrics of the ER-based model fusion with the other three models. Table 1 showed that the ER-based model fusion performed the best in terms of accuracy, recall

rate, and AUC, while its precision index was slightly lower than GBDT and neural network models. Overall, the ER-based model fusion performed better than the other three models. Therefore, the proposed two-stage approach demonstrated its effectiveness in improving the performance of using multiple machine learning models.

## **5. Conclusions and Implications**

This study proposed and illustrated a two-stage machine learning approach for modeling the CLV of airline passengers. This study contributes to the CLV literature by integrating the internal data and external data into CLV models and introducing the ER-based model fusion approach to improve prediction accuracy. This study provides an important methodological contribution to accessing CLV from a comprehensive perspective. Because the existing CLV models only utilize the company's internal data, the existing models are biased and cannot evaluate the passenger's overall value to the aviation industry. For example, Passenger A flew with the Airlines X once in 2019 and spent 100 USD, while he spent over \$10,000 USD in other flights with other airlines. Apparently, Passenger A has high buying power, and he seems not to prefer Airline X over other airlines. If merely using the internal data, the Airlines X would consider him as a low-value customer and mistakenly neglect his potential to bring a higher value. However, Passenger A is an important customer who needs more marketing efforts, because the person may change his brand preference and purchase more tickets if the Airlines X adopts appropriate marketing strategies to improve the customer relationship with Passenger A. Therefore, by introducing the data from TravelSky, this study overcomes the limitation of the existing CLV models and enables the airline companies to identify those high-value passengers from the perspective of whole aviation network.

We illustrated the proposed two-stage approach with a case study of a Chinese airline company. The results indicated that the ER-based model fusion performed better than other machine learning

models in the classification of high-value passengers. This study contributes to the ER literature by innovatively applying the ER method to fuse other traditional models of CLV assessment. Although ER has been applied in many studies, our study is the first attempt to apply the ER method for assessing CLV (Yang, Xu, Xie, & Maddulapalli, 2011; Wang, Tang, & Wang, 2014). As demonstrated in the case study, the proposed two-stage approach could help airline companies to identify high-value passengers more accurately, so that airlines can develop personalized CRM strategies to improve and maintain customer relationships more wisely. Given the current competition among airline companies are very intensive, this study offers very important managerial implications for airlines.

One of the possible limitations of this study is that we only conducted binary classification in the case study. For future studies, we suggest multiple-class classification should be further investigated to help the airlines to distinguish passengers' CLV levels (e.g., premium, high, medium, and low). This study combined the airlines' internal data and external data from TravelSky, however, we do not incorporate external data from other third-parties such as passengers' consumption at airport duty-free stores and passengers' data on social media platforms. Future research should further explore external data related to CLV and incorporate them into the assessment model of CLV.

## **References**

- Araujo, L., & Kjellberg, H. (2015). Forming cognitions by investing in a form: Frequent Flyer Programs in US air travel post-deregulation (1981–1991). *Industrial Marketing Management*, 48, 68-78.
- Armitage, D. W., & Ober, H. K. (2010). A comparison of supervised learning techniques in the classification of bat echolocation calls. *Ecological Informatics*, 5(6), 465-473.
- Atkinson, T. M., Sit, L., Mendoza, T. R., Fruscione, M., Lavene, D., Shaw, M., ... & Basch, E. M. (2010). Confirmatory factor analysis to evaluate construct validity of the Brief Pain Inventory (BPI). *Journal of Clinical Oncology*, 28(15\_suppl), e19505-e19505.



- Castéran, H., Meyer-Waarden, L., & Reinartz, W. (2017). Modeling customer lifetime value, retention, and churn. *Handbook of Market Research*, 1-33.
- Dedu, V., & Ganea, T. A. (2009). Developing a rating model on a statistical basis. *Theoretical and Applied Economics*, 1(1), 33.
- Dempster, A. P. (2008). Upper and lower probabilities induced by a multivalued mapping. In *Classic works of the Dempster-Shafer theory of belief functions* (pp. 57-72). Springer, Berlin, Heidelberg.
- Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., ... & Sriram, S. (2006). Modeling customer lifetime value. *Journal of service research*, 9(2), 139-155.
- Gupta, S., Lehmann, D. R., & Stuart, J. A. (2003). Valuing Customers. *SSRN Electronic Journal*.
- Hughes, A. M. (1996). *Boosting response with RFM. Marketing Tools*, 4-8.
- Jain, D., & Singh, S. S. (2002). Customer lifetime value research in marketing: A review and future directions. *Journal of interactive marketing*, 16(2), 34.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.
- Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, 160, 3-24.
- Kullback, S. (1997). *Information theory and statistics*. Courier Corporation.
- Mahmud, T., Rahman, K. N., & Hossain, M. S. (2013). Evaluation of job offers using the evidential reasoning approach. *Global Journal of Computer Science and Technology*.
- Pfeifer, P. E., Haskins, M. E., & Conroy, R. M. (2005). Customer lifetime value, customer profitability, and the treatment of acquisition spending. *Journal of managerial issues*, 11-25.

- Rahman, M. S., Mahmud, M. A., Pota, H. R., & Hossain, M. J. (2014). Distributed multi-agent scheme for reactive power management with renewable energy. *Energy conversion and management*, 88, 573-581.
- Schmittlein, D. C., Morrison, D. G., & Colombo, R. (1987). Counting your customers: Who-are they and what will they do next?. *Management science*, 33(1), 1-24.
- Schuld, M., Sinayskiy, I., & Petruccione, F. (2016). Prediction by linear regression on a quantum computer. *Physical Review A*, 94(2), 022342.
- Stone, M., Woodcock, N., & Wilson, M. (1996). Managing the change from marketing planning to customer relationship management. *Long Range Planning*, 29(5), 675-683.
- Si, X. S., Hu, C. H., Yang, J. B., & Zhou, Z. J. (2011). A new prediction model based on belief rule base for system's behavior prediction. *IEEE Transactions on Fuzzy Systems*, 19(4), 636-651.
- Tang, O., & Musa, S. N. (2011). Identifying risk issues and research advancements in supply chain risk management. *International journal of production economics*, 133(1), 25-34.
- Venkatesan, R., & Kumar, V. (2004). A customer lifetime value framework for customer selection and resource allocation strategy. *Journal of marketing*, 68(4), 106-125.
- Vinod, B. (2016). Big Data in the travel marketplace. *Journal of Revenue and Pricing Management*, 15(5), 352-359.
- Wang, Y. M., Yang, J. B., Xu, D. L., & Chin, K. S. (2006). The evidential reasoning approach for multiple attribute decision analysis using interval belief degrees. *European Journal of Operational Research*, 175(1), 35-66.
- Wang, Z. L., Tang, K., & Wang, X. (2014). A multi-model soft sensing method based on DS and ARIMA model. *Control and Decision*, 29(7), 1160-1166.

- Yang, J. B., & Singh, M. G. (1994). An evidential reasoning approach for multiple-attribute decision making with uncertainty. *IEEE Transactions on systems, Man, and Cybernetics*, 24(1), 1-18.
- Yang, J. B., Xu, D. L., Xie, X., & Maddulapalli, A. K. (2011). Multicriteria evidential reasoning decision modelling and analysis—prioritizing voices of customer. *Journal of the Operational Research Society*, 62(9), 1638-1654.
- Yang, Y., Xu, D. L., Yang, J. B., & Chen, Y. W. (2018). An evidential reasoning-based decision support system for handling customer complaints in mobile telecommunications. *Knowledge-Based Systems*, 162, 202-210.
- Yokota, F., & Thompson, K. M. (2004). Value of information analysis in environmental health risk management decisions: past, present, and future. *Risk analysis: an international journal*, 24(3), 635-650.
- Yoo, S., & Hanssens, D. M. (2005). Modeling the sales and customer equity effects of the marketing mix. *UCLA Anderson School of Management*, 2, 1-42.