Estimating Customer Lifetime Value Using Machine Learning Techniques

DOI:
10.5772/intechopen.76990

Document Version
Final published version

Link to publication record in Manchester Research Explorer

Citation for published version (APA):
Chen, S. (2018). Estimating Customer Lifetime Value Using Machine Learning Techniques. In C. T
Data Mining (pp. 17-34). IntechOpen. https://doi.org/10.5772/intechopen.76990

Published in:
Data Mining

Chapter 2

Estimating Customer Lifetime Value Using Machine Learning Techniques

Sien Chen

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.76990

Abstract

With the rapid development of civil aviation industry, high-quality customer resources have become a significant way to measure the competitiveness of the civil aviation industry. It is well known that the competition for high-value customers has become the core of airline profits. The research of airline customer lifetime value can help airlines identify high-value, medium-value and low-value travellers. What is more, the airline company can make resource allocation more rational, with the least resource investment for maximum profit return. However, the models that are used to calculate the value of customer life value remain controversial, and how to design a model that applies to airline company still needs to be explored. In the paper, the author proposed the optimised China Eastern Airlines passenger network value assessment model and examined its fitting degree with the TravelSky value score. Besides, the author combines China Eastern Airlines passenger network value assessment model score with loss model score to help airlines find their significant customers.

Keywords: customer lifetime value, estimating, machine learning

1. Introduction

In the context of customer relationship management, customer lifetime value (CLV) or customer equity (CE) becomes important because it is a disaggregate metric to evaluate marketing decisions [1], which can be utilised to allocate resources appropriately and identify profitable consumers [2]. Companies are looking forward to better approaches to create value and optimise their market offerings to appeal to customers and make profits [3]. Many firms are utilising CLV regularly to control and supervise the strategies of marketing as well as evaluate the business success. For companies, it is of interest to know how much net benefit it can expect from their customers. It is recognised that clv has become a significant component of companies' central strategy [4, 5]. CLV of customers at present and in the future can be a good proxy of the general corporate value [6]. Meanwhile, at each point in each customer's lifetime with the firm, the firm would like to form some expectation regarding the lifetime value of that customer.

2. Definition of CLV

Customer valuation has been discussed by several papers in the customer relationship management literature, for example, Dwyer [7], Berger and Nasr [8], Rust et al. [9] and Blattberg and Malthouse [10].

Dwyer [7] and Berger and Nasr [8] firstly provided a framework using the lifetime value of a customer. Then Gupta and his colleagues [6] found that the earnings of a company, and hence its value, are a function of the total customer lifetime value (CLV), defined as the discounted value of the future profits yielded by customers to the company, in other words, the value of a customer as the expected sum of discounted future earnings, where a customer generates a profit margin for each period. Moreover, a customer lifetime value (CLV) stands for the expected benefits' current value [7] and the equity of customer approaches to marketing [11, 12]. And CLV plays a major role in the marketing of the relationship [13]. The relationship with customers in the relationship marketing can be considered as the capital assets that require proper management [14].

3. Related work

In measuring customer lifetime value, a standard approach is to estimate the present value of the net benefit to the firm from the customer over time [1]. Researchers have suggested various methods to use customer-level data to measure the CLV [8, 9, 15–17]. However, the relationship between customer purchase behaviour and customer lifetime is not specific [15–19], if firms observed the customer defections, and longer customer lifetime implies higher customer lifetime value [20–22]. Different models for measuring CLV are different at estimates of the expectations of future customer purchase behaviour.

3.1. Methods of CLV prediction

3.1.1. CLV model

CLV is typically defined and estimated at an individual customer or segment level. This allows us to differentiate between customers who are more profitable than others rather than merely examining average profitability. The issue is to predict the future profits when the timing and the benefit of future transactions are unknown as discussed in Mulhern [23] and Bell et al. [24]. It is proposed by Gupta and other scholars [25] that CLV for a customer is [6, 19]:

$$\mathrm{CLV} = \sum\nolimits_{t=0}^{T} \frac{(P_t - C_t)r_t}{(1+i)^t} - AC \tag{1}$$

It is proposed by Gupta and other scholars [18] that CLV for a customer is [19, 36]:

where:

= price paid by a consumer at time t.

= direct cost of servicing the customer at time t.

= discount rate or cost of capital for the firm.

= probability of customer repeat buying or being 'alive' at time t.

AC = acquisition cost.

T = time horizon for estimating CLV.

Another review of CLV model sees Jain and Singh [26]. Linear regression with the variance that stabilises the transformation forecasted with the ordinary least square is the first approach. Selecting a stable variance transformation can be informed by residual plots [27]. As shown by Neter et al. [28], the linear regression forecasted with iteratively reweighted least square is the second approach of regression. IRLS is another means to solve the heteroscedasticity issue.

3.1.2. RFM model

For the sake of simplicity, the only predictor variables in these models are the recency, frequency and monetary (RFM) type, Buckinx and Van den Poel [29], and the variables of RFM are sound predictors for CLV [15, 16].

The models of RFM have been utilised in direct marketing for three decades developed to target marketing programmes at specific customers with the objective to improve response rates. Studies show that customers' response rates vary the most by their recency, followed by their purchase frequency and monetary value [30]. Before these models, companies typically used demographic profiles of customers for targeting purposes. However, research strongly suggests that past purchases of consumers are better predictors of their future purchase behaviour than demographics.

They have many restrictions though RFM, or other models of scoring try to forecast customers' behaviour in the future and are therefore associated with CLV implicitly [15, 16, 31]. Firstly, in the next periods, the behaviour can be predicted by the models. However, to estimate CLV, we need to estimate customers' purchase behaviour not only in Period 2 but also in Periods 3, 4, 5 and so on. Secondly, RFM variables are real underlying behaviour's imperfect index stemmed from a real distribution. The models of RFM have neglected this part. Thirdly, the previous behaviour of customers can be an outcome of the company's previous marketing promotion, which has been ignored by the models. In spite of the restrictions, due to the implementation in real practice, the models of RFM are the core of the industry.

One fundamental limitation of RFM models is that they are scoring models and do not explicitly provide a number for customer value. However, RFM is essential past purchase variables that should be good predictors of future purchase behaviour of customers. Fader et al. [15, 16] showed how RFM variables could be used to build a CLV model that overcomes many of its limitations.

3.1.3. NBD-Pareto model

A popular method is the negative binomial distribution (NBD)-Pareto model introduced by Schmittlein et al. [32], which is referred by several authors [23, 26, 33] as a powerful technique to provide the situation where past customer purchase behaviour is used to predict the future probability of a customer remaining in business with the firm.

To forecast the CLV and integrate the transaction profits, some adoptions are conducted as the model of NBD-Pareto estimates the activity probability and the transaction number of a customer. Made by the NBD-Pareto for the forecast, an essential assumption refers to the independence between the relevant profit for every transaction and the transaction number of a customer. According to the prediction of a majority of papers, a two-step scheme to CLV modelling is being utilised by CLV [16, 17, 34]. Firstly, the transaction number of every person in the future will be forecasted. Subsequently, the mean profit for every transaction can be forecasted. At the level of customers, the values can be predicted. It generates a CLV approximation for every customer if the future transaction number and the mean profit for every transaction can be concluded.

In Fader and Hardie [15, 16], the maximum likelihood estimation (MLE) for an individual with purchase history is shown to describe the NBD-Pareto submodel. Utilising the approach of moments is an alternative to the MLE. However, similar results can be generated [19]. A person can forecast the transaction number that will be made by a customer in the future or predict the possibility for him or her to be alive when the parameters can be forecasted. As discussed by Schmittlein and Peterson [17], in the situation where customer lifetimes are observed, the NBD-Pareto model has limitations and is not suitable.

Another approach that can naturally incorporate past behavioural outcomes into future expectations is a Bayesian approach [35]. Bayesian approaches could integrate the previous data and information into the model's structure via the prior distribution of the CLV drivers.

3.1.4. Computer science models

The vast computer science literature in data mining, machine learning and nonparametric statistics has generated many approaches that emphasise predictive ability. These include projection-pursuit models; neural network models [36]; decision tree models; spline-based models such as generalised additive models, multivariate adaptive regression splines and classification and regression trees; and support vector machines. Lots of the methods might be more applicable to the research on the value in customers' lifetime.

In a recent study, Cui and Curry [37] conducted extensive Monte Carlo simulations to compare predictions based on multinomial logit model. Besides, Giuffrida et al. [38] reported that a multivariate decision tree induction algorithm outperformed a logit model in identifying the best customer targets for cross-selling purposes.

Due to the high focus that academics in marketing emphasise on interpretability and a parametric setup, these approaches remain little known in the marketing literature. However, given the importance of prediction in CLV, these methods need a closer look at the future.

3.1.5. RFMc model

The meaning of individual passenger value is calculating the traveller's particular value for the airline company based on the passenger's consumption data. It also refers to the passengers' profit contribution to the airline company.

Based on the characteristics of civil aviation, the fare discounts corresponding to class C are introduced to represent the level of value which passenger's consumption contributes to airlines. The RFMc model is proposed to calculate the civil aviation passengers' individual value, where R is the closeness coefficient of flight time, F is the total number of flights in a period of time and Mc is the passengers' relative total amount of flights calculated with the class of flight.

(1) Mc: the passengers' relative total amount of flights.

Calculate the total amount of relative consumer consumption Mc based on the fare weight of class c (corresponding fare discount); see formula (2):

$$M_c = \sum_{i=1}^{k} m_i * c_i \tag{2}$$

In the formula (2), c_i represents the fare discount on the traveller's ith flight, m_i is the fare of the traveller's ith plane, and k is the number of tickets purchased.

(2) R: the approximate coefficient of flight time.

The latest flight time t: the interval between the last flight time and the current time (the time when using the model to calculate the passenger's value).

The average turnaround time of flight t_0: the average of the two adjacent flights' time interval; see formula (3):

$$t_0 = \begin{cases} \sum_{i=1}^{n-1} t_i/(n-1) & n > 1 \\ t_s & n = 1 \end{cases} \tag{3}$$

In the formula (3), n is the gross number of passenger flights, t_i is the passenger's flight time interval between ith and (i + 1)th, and t_s is the average turnaround time of the precalculated whole passenger set.

The approximate coefficient of flight time R: the possibility that passengers take the plane again; see Eq. (4):

$$R = \begin{cases} 1 & t \leq t_0 \\ \frac{t_0}{t} & t > t_0 \end{cases} \tag{4}$$

The average flight turnaround time t_0 reflects the expectation of the interval between passengers' two contiguous flights. As the latest flight time t is less than or equal to the average turnaround time t_0, the value of R is 1; when t is greater than t_0, the possibility of passengers taking off again is gradually reduced, and R is slowly decreased.

(3) F: the passengers' flight frequency.

The passengers' flight frequency F reflects the activity and loyalty of passengers. It is acknowledged that the activity and loyalty affect the CLV to the airline company. The greater the take-off frequency, the higher the activity and loyalty degree, which can lead to the greater passenger's value to the airline. In general, the passengers' relative total amount of flights, the approximate coefficient of flight time and the passengers' flight frequency weighted sum, to obtain the passengers' value 'v', see Eq. (5):

$$v = \omega_1 M_c + \omega_2 R + \omega_3 F \tag{5}$$

In formula (5), ω_1, ω_2 and ω_3 are each indicator's weight coefficients. Considering the different measurement of different indicators, Mc, R and F should be standardised and then weighted summation.

3.1.6. MRE model

Passenger co-occurrence relationship includes the same order explicit co-occurrence relationship and different orders implicit coordination relationship. MRE multi-relational evaluation model combines order data and departure data, quantifies the explicit and implicit relationship between passengers and integrates time to make the comprehensive multi-relational evaluation.

(1) The same order co-occurrence relationship.

The same order co-occurrence relationship refers to the passenger relationship in the same order. The passenger's the same order relationship includes the number of passengers in the order, the difference between passenger class and order generation date. Based on PNR data to establish the whole passengers' same order relationship, use P_{ij} to show the sequence of the same order relationship between passenger i and passenger j. P_{ij} [k] = [∣ c_i [k] -c_j [k] ∣, s [k], t_p [k]] is the kth record in the sequence, which indicates the data from the passenger i and passenger j's same order, where s [k] is the number of passengers of the order, t_p [k] is the order generation date and c_i [k] is the class of passenger i in the order (corresponding to the fare discount).

According to the sequence of the passenger's same order relation, passenger's same order relationship score is calculated. P'_{ij} shows the total score of the same order relationship between passenger i and passenger j; see formula (6):

$$\mathrm{p}'_{\mathrm{ij}} = \Sigma_k s_P[k] = \Sigma_k \frac{1}{\sqrt{\mathrm{s}[k] \times \left(\left|c_i[k] - c_j[k]\right| + 1\right)}} \tag{6}$$

In the formula (6), $s_p[k]$ is the score of the kth same order between passenger i and passenger j.

(2) Passenger company relationship

Company relationship: company relationship is defined by the author as the passenger-company relationships on the same flight which include coincidental company and appointed company. A co-occurrence relationship includes the date of flight departure, passenger seat distance, check-in sequence number distance, class rank difference and other attributes. According to the whole passengers' company relationship based on the departure data, D_{ij} is denoted as the sequence of company relationship between passenger i and passenger j. D_{ij} [k] = [| d_{ci} [k] |, | d_{seat} [k] |, | d_{class} [k], t_d [k]] is the kth record in the sequence, which represents the kth flight data of passenger i and passenger j when they fly together. Among these, t_d [k] is the flight departure date, d_{ci} [k] represents the check-in distance between passenger i and passenger j, d_{seat} [k] represents the Euclidean distance between passenger i and passenger j's flight seats and d_{class} [k] represents the class difference between passenger i and passenger j. According to the processed sequence of passenger-company relationship, the passenger-company relationship score can be calculated, where D'_{ij} is used to show the total company relationship score of passenger i and passenger j, and the formula is given as

$$D'_{ij} = \sum_k S_d[k] \tag{7}$$

$$S_d[k] = \frac{\omega_1}{d_{ci}[k]} + \frac{\omega_2}{d_{seat}[k]} + \frac{\omega_3}{d_{class}[k] + 1} \tag{8}$$

In formulas (7) and (8), ω_1, ω_2 and ω_3 are the impact factors of check-in sequence number distance, seat distance and class difference on passenger-company relationship score. S_d [k] is the kth company relationship score between passenger i and traveller j.

(3) Time involved multi-relational comprehensive evaluation

Passenger value is unevenly distributed according to the edge weight. The scientific and accurate calculation of the edge weight directly affects the result of passenger value for the reason that the closer the passenger relationship is, the higher the value distributed. The RFM model predicts the possibility of customer repurchasing on the basis of customer consumption proximity R. Similarly, we also think that the civil aviation-passenger relationship is also connected with time: The passengers that fly together in the last few days are more likely to travel together again and have a closer relationship. In contrast, even if they have been together for many times, but no record of company in the past 2 years, we also have to consider whether the passenger relationship has disappeared. Due to the above considerations, we set the observation time window to observe the passenger relationship and bring in the time attenuation factor τ to make the passenger's relationship time perceptive. Assuming that the same last order (or same flight) of traveller i and traveller j is t, the time attenuation factor τ of the same order (or company) relationship between passenger i and passenger j can be expressed as

$$\tau = \frac{t - t'}{T - t'} \tag{9}$$

where **T–t** 'is the length of the observation time window, **T** is the end time of the time window and **t'** is the beginning time of the time window. If t ≤ t 'means that the passenger does not have the same order (or company) relationship in the observation time window, then the relationship is considered to disappear, and assume that $\tau = 0$. After introducing the time attenuation factor, the score of the same order passenger relationship can be expressed as formula (10), and the score of passenger-company relationship can be expressed as formula (11):

$$P'_{ij} = \tau_{pij} \times \sum_k S_p[k] \tag{10}$$

$$D'_{ij} = \tau_{Dij} \times \sum_k S_d[k]$$

where τ_{Pij} is the time attenuation factor of the passenger i and the passenger j's same order relationship and τ_{Dij} is the time attenuation factor of the same order relationship between the passenger i and the passenger j.

Standardise the passengers' company relationship score and the same order relationship score, and then weight and sum to get the total passenger relationship score. The formula is

$$\mathrm{W}_{ij} = \omega_{\mathrm{P}} P'_{ij} + \omega_{\mathrm{d}} D'_{ij} \tag{11}$$

where W_{ij} represents the total score of the relationship between passenger i and passenger j, ω_p, ω_d, followed by the same order relationship weight and company relationship weight, $\omega_p < \omega_d$.

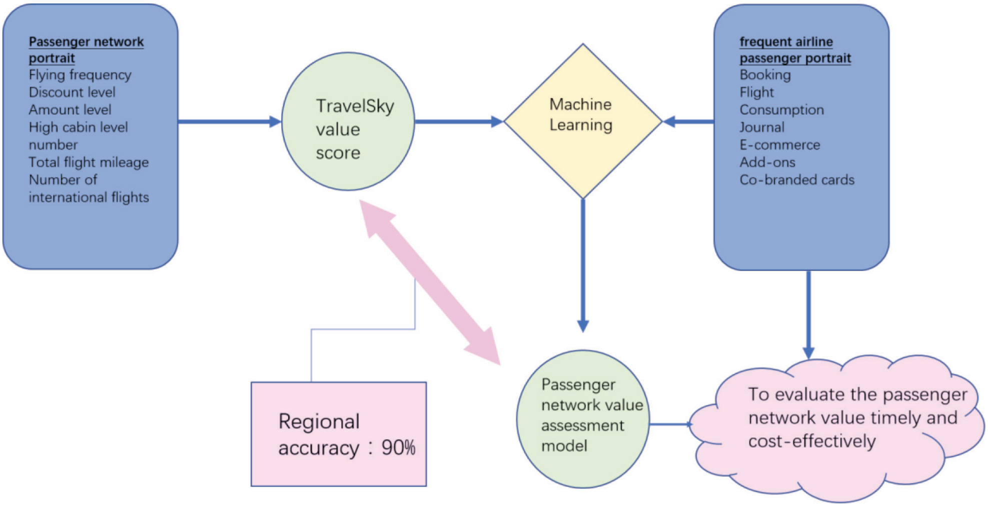
3.2. CLV prediction accuracy

Fit is the criterion suggested in the data-mining literature [39–41] for problems where the primary objective is making predictions that are as accurate as possible. As measures of prediction accuracy, Glady et al. [42] used the mean absolute error (MAE) and root mean square error (RMSE) between the actual value and the forecast of value in customers' lifetime. The 1% trimming can be used for the MAE and RMSE to enhance the strength to potential outliers in the set of data.

4. Passenger network value assessment model

4.1. Model description

Based on the dimensions of flying frequency, discount level, amount level, total flight mileage and number of international flights, etc. in the past year, TravelSky makes a comprehensive assessment on the value of passengers every month and form a scale of 0–100 value score. Which is called TravelSky Value Score. Passenger network value assessment based on the internal data of China Eastern Airlines, using airline frequent personal attributes and the airline's internal flight network's behaviour to estimate the TravelSky Value Score by using the advanced machine learning model. By fitting the TravelSky Value Score to the XGBoost model, a high fitting accuracy rate can be obtained, therefore helping the airline to evaluate the passenger network value timely and cost-effectively and to provide follow-up passenger segmentation and precision marketing services.



4.2. Data collection: frequent airline passenger portrait

First, collect data from relational database, log system, file system, document, picture, video, voice and other sources of different formats; analyse and identify the data. Then, focus on the business to identify and comprehend the information from the data. After that, extract valuable data fusion to the data platform. The dimension of frequent airline passenger portrait involves more than 300 variables including booking, flight, consumption, journal, e-commerce, add-ons and co-branded cards.

4.3. Passenger network value assessment model

4.3.1. The inputs of the model

The input of the model is regarded as the relevant data or information which is used for computer processing. More specifically, in the process of the model application, input refers to the data of human and human behavioural characteristics. In the case of China Eastern Airlines, the inputs of passenger network value assessment model include 300+ variables, such as member current level, the highest consumption points in the last 3 months, the average delay time, how much changes of the air ticket endorsement, etc. However in general, the 300+ variables can be categorised into booking, flight, consumption, journal, e-commerce, add-ons and co-branded cards.

4.3.2. The outputs of the model

The outputs of CEA passenger network value assessment model is estimated CEA passenger value score.

4.3.3. The mechanism of the model

XGBoost is adopted as the mechanism in the paper.

(1) The introduction of XGBoost

XGBoost is a scalable machine learning system for tree boosting. The system is accessibly regarded as an open source package2.

XGBoost most prominent feature is that it can automatically use the CPU's multithreaded parallel while improving the algorithm to enhance the accuracy. Its debut is the Kaggle Higgs Sub Sign Recognition Contest, because of its superior efficiency and high predictive accuracy and it caught the attention of contestants in the competition forum.

(2) The Objective function of the optimisation model is

$$Ob_j(\theta) = L(\theta) + \Omega(\theta) \tag{12}$$

where $L(\theta)$ is error function which proves how well our model fits the data. $\Omega(\theta)$ is regularisation term, which is used to punish complex models [43].

The error function encourages the optimisation model to fit the training data, while the regularisation term helps the simpler model. Because when the model is simple, the randomness of the fitting degree of the finite data is relatively small and is not accessible to overfitting, making the prediction of the final model more stable.

The optimisation objective function in this case is

$$Ob_j(\theta) = \sum_i^n l\left(y_i, \widehat{y}_i\right) + \sum_{k=1}^K \Omega\left(f_k\right) \tag{13}$$

In this function, $\widehat{y}_i$ is estimated passenger network value score and, $\widehat{y}_i$ is TravelSky value score.

For more objective function derivation process, please refer to 《XGBoost: A Scalable Tree Boosting System》.

4.4. The performance of passenger network value assessment model

4.4.1. Model main parameters

Tree depth, 6; step size, 0.1; maximum number of iterations, 66.

4.4.2. Model indicators

rmse: Training Set 11.9455 and Test Set 13.02934.

R^2, 0.3939464.

4.4.3. The model main feature variables

Feature		**Gain**	**Cover**	**Frequency**
wd_tk_bef_mean_hur_curr	Average advance booking time (next time window) (hours)	**0.186679089**	0.055902665	0.013308573
wd_tk_bk_next_nbr_curr	The number of future booking (the next time window)	**0.145693046**	0.049101542	0.013618075
wd_tk_bk_mean_intv_curr	Average booking time interval (next time window)	**0.050121893**	0.05435901	0.023522129
travel_max_intv_3m	The maximum flight time interval in the last 3 months	**0.045163104**	0.019874175	0.007118539
wd_tk_bk_mean_intv_prev	Average booking time interval (last time window)	**0.028847336**	0.008726862	0.013308573
ap_income_channel_3m.非航累积	Accumulate the main channel of the last 3 months: non-flight accumulation	**0.023556519**	0.015727313	0.008047044
ap_income_sum_1y	The total number of points accumulated in the most recent year	**0.018971991**	0.028146008	0.010523058
ticket_bef_max_intv_1y	The largest number of days in advance tickets in the latest year	**0.014451596**	0.030321255	0.009285051
wd_tr_zj_curr	Average early check-in time (the next time window)	**0.01381261**	0.02388794	0.006499536
wd_tk_bk_nbr_curr	Booking times (next time window)	**0.013618875**	0.01638353	0.006499536

4.4.4. Distribution of TravelSky value and forecast value

Separately observe their scores, and it can be seen that the scores are all concentrated in the high segment. In particular, 63.13% of the passengers get 100 TravelSky Value scores.

Summation items: the number of people	
TravelSky value	**Summary**
0	1.04%
91	0.36%
92	0.44%
93	0.61%
94	0.82%
95	1.17%
96	1.82%
97	3.14%
98	6.35%
99	16.56%
100	63.13%
Total	**100.00%**

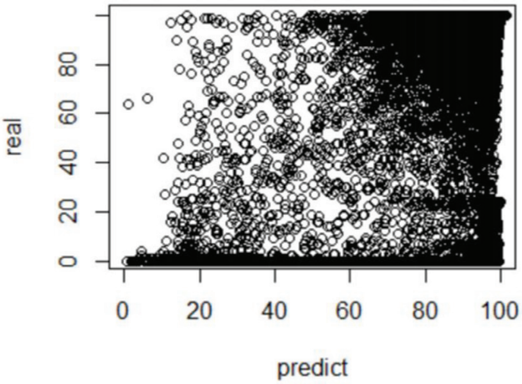
Summation items: the number of people	
forecast value **(rounding)**	**Summary**
0	0.00%
91	1.08%
92	1.40%
93	1.94%
94	3.03%
95	4.58%
96	6.66%
97	15.69%
98	49.32%
99	9.08%
100	0.07%
Total	**100.00%**

4.5. Model evaluation report: TravelSky value fit report

Using more than 300 features of CEA loss model and 240,000 passenger data of loss model, the TravelSky value score is fitted to the Xgboost model [44, 45].

4.5.1. Cross-contrast the TravelSky value score and the forecast value

Cross-contrast the TravelSky value score with the forecast value; visualise the data and present it in the form of the charts below.



PivotTable: The horizontal axis represents the 10-point range where the value score fits with CEA data. For example, 1 indicates [0, 10], 2 indicates [11, 20], and similarly, 10 indicates

(90,100). The vertical axis represents the 10-point interval in which the avionics value score is located.

Summation items: the number of people	**Predicted value (divided by 10 and rounded)**										
TravelSky value **(divided by 10 and rounded)**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Total**
1	0.04%	0.25%	0.14%	0.09%	0.08%	0.08%	0.11%	0.14%	0.23%	0.68%	1.84%
2	0.00%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.15%	0.26%
3	0.00%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.04%	0.24%	0.37%
4	0.00%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.03%	0.05%	0.16%
5	0.00%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.03%	0.04%	0.07%	0.21%
6	0.00%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.03%	0.06%	0.10%	0.26%
7	0.00%	0.00%	0.01%	0.01%	0.01%	0.02%	0.02%	0.04%	0.08%	0.15%	0.34%
8	0.00%	0.01%	0.01%	0.01%	0.01%	0.02%	0.03%	0.08%	0.13%	0.33%	0.62%
9	0.00%	0.01%	0.01%	0.01%	0.01%	0.03%	0.06%	0.12%	0.34%	0.96%	1.56%
10	0.00%	0.01%	0.01%	0.02%	0.03%	0.07%	0.19%	0.62%	3.33%	90.11%	94.39%
Total	**0.04%**	**0.32%**	**0.23%**	**0.18%**	**0.20%**	**0.27%**	**0.50%**	**1.11%**	**4.29%**	**92.85%**	**100.00%**

Evaluation criteria:

1. Calculate the accuracy of 10-point interval: 90.64%.
2. As the forecast scores are mainly concentrated around 98 points, the proportion of people between 0 and 90 points is relatively low which belongs to the low-value area. Therefore, the author will divide '0–90 points' into a category. Using the 400,000 senior frequent passengers of China Eastern Airlines's portrait features to fit the TravelSky value score, the accuracy of the evaluation is up to 90% with ten-point interval. 92.85% of the passenger network value assessment model (CEA model) is located in the 91–100 value range. Ninety-seven percent (90.11%/92.85% = 97%) of the TravelSky value score is also located in the [91,100] value range (as shown in the following table).

The proportion of the population		**CEA value**		**Total**
		[0, 90]	**[91, 100]**	
Travel Sky	**[0, 90]**	2.87%	2.74%	5.61%
	[91, 100]	4.28%	90.11%	94.39%
Total		7.15%	92.85%	100.00%

4.6. Module application

Based on the accuracy of the passenger network value assessment model combined with the prediction of passenger loss probability in the next 6 months, it is necessary to give priority to reach the target of 'high network value and high risk of loss in the next 6 months' passenger groups. Thus, it can help marketing accurate positioning.

4.6.1. Passenger loss model

A. Definition of loss: The number of flight phase in the next 6 months is at least decreasing ten absolute flight phases or reducing 50%.

B. The loss model Xgboost main features.
Summary (model).

Importance of features in the XGBoost model.

Feature		Gain	Cover	Frequency
wd_tr_12h_dep_curr	Delay [1, 2] Number of flight phase (next time window)	0.085744211	0.060886127	0.017437145
wd_tr_24h_dep_series_c	Delay [2,4] Number of flight phase (how many changes)	0.064961789	0.032607435	0.010948905
wd_tr_24h_dep_series_b	Delay [2,4] Number of flight phase (whether changed)	0.054268923	0.040321268	0.01216545
deploy_arr_mean_tm_3m	Average delay of flight arrival in the last 3 months (in minutes)	0.048203621	0.034246522	0.016626115
wd_tr_y_nbr_curr	Economy class travel flight phase number (next time window)	0.044511041	0.057238477	0.01297648
travel_f_max_intv_3m	The latest 3-month maximum flying time interval	0.040740831	0.047666125	0.01865369
deploy_1h_nbr_3m	The number of flight phases which flight delays of 1 hour in the last 3 months	0.018180959	0.014574099	0.01540957
wd_pt_aft_lvl_labels_b.银卡	Member level (end) (from A to B): silver card	0.016429216	0.017211579	0.004460665
deploy_arr_mean_tm_1y	Average delay of flight arrival in the most recent year (in minutes)	0.015352698	0.008028182	0.00729927
wd_tk_bk_next_nbr_prev	The number of future booking (the last time window)	0.014210402	0.010862506	0.01054339
y_hd_cnt_3m	The last 3-month economy class flight phase number	0.011647629	0.01455673	0.01865369
wd_tr_dpt_mean_dep_curr	Average delay time (departure, minute) (next time window)	0.010571347	0.01455938	0.00486618
wd_pt_aft_lvl_labels_b.小飞人	Member level (end) (from A to B): supermaster	0.010329831	0.011708833	0.0162206
deploy_dpt_mean_tm_3m	Average take-off delay in the last 3 months of flight (in minutes)	0.01016305	0.013088575	0.01540957
y_hd_cnt_6m	The last 6-month economy class travel flight phase number	0.010010692	0.004976274	0.01459854
wd_tk_sum_amt_curr	Total booking amount (next time window)	0.009631331	0.003969202	0.0081103
travel_f_mean_intv_1y		0.009385889	0.020736252	0.011759935

Feature		Gain	Cover	Frequency
	First-class average flying interval in the latest year			
deploy_dpt_max_tm_3m	The maximum time of take-off delay in the last 3 months (unit: minutes)	0.009230547	0.007903106	0.0081103
y_hd_cnt_1y	The number of economy class flight phase in the latest year	0.008812939	0.003531476	0.011759935
wd_tr_f_nbr_curr	First-class flight phase number (next time window)	0.008393582	0.019247921	0.01135442

The resulted model fits the entire dataset, and the relative importance of each variate can be viewed by importance_xgb () or simpler summary () as above.

4.6.2. Loss model score combines the forecast value score to select the key population

Summation items: the number of people	**Estimated loss rate (10% interval)**										
Forecast value (divided by ten and rounded)	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Total**
1	0.00%	0.00%	0.01%	0.01%	0.01%	0.01%	0.00%	0.00%	0.00%	0.00%	0.04%
2	0.00%	0.04%	0.07%	0.09%	0.06%	0.04%	0.02%	0.00%	0.00%	0.00%	0.32%
3	0.00%	0.03%	0.04%	0.05%	0.05%	0.03%	0.01%	0.01%	0.00%	0.00%	0.23%
4	0.00%	0.02%	0.04%	0.04%	0.03%	0.03%	0.01%	0.00%	0.00%	0.00%	0.18%
5	0.00%	0.02%	0.03%	0.04%	0.04%	0.03%	0.02%	0.01%	0.00%	0.00%	0.20%
6	0.00%	0.02%	0.05%	0.06%	0.05%	0.05%	0.03%	0.01%	0.00%	0.00%	0.27%
7	0.00%	0.04%	0.10%	0.11%	0.10%	0.08%	0.04%	0.02%	0.01%	0.00%	0.50%
8	0.00%	0.08%	0.21%	0.25%	0.23%	0.17%	0.11%	0.04%	0.01%	0.00%	1.11%
9	0.01%	0.31%	0.73%	0.98%	0.93%	0.71%	0.41%	0.17%	0.04%	0.01%	4.29%
10	0.24%	4.26%	12.40%	18.59%	20.64%	18.18%	12.02%	5.32%	1.13%	0.07%	92.85%
Total	**0.26%**	**4.82%**	**13.68%**	**20.24%**	**22.14%**	**19.31%**	**12.67%**	**5.59%**	**1.21%**	**0.08%**	**100.00%**

An orange group means that both high loss scores (high likelihood of loss in the next 6 months) and high-value scores (up to 90 points) fit well with TravelSky value score.

5. Conclusion

In this paper, the author first described the definition of customer lifetime value (CLV) and demonstrated the approach to estimating customer lifetime value by proposing various customer lifetime value models and illustrating the criterion to predict customer lifetime value accuracy. The aim is to provide the theoretical basis for the airline customer lifetime value estimation research. After that, a numeral case of China Eastern Airlines was given to show the practicability and veracity of China Eastern Airlines passenger network value assessment model with assessing their fitting accuracy rate with the TravelSky value score. The ambition is combining forecast value score calculated by China Eastern Airlines passenger network value assessment model with loss model score to select the critical population.

Author details

Sien Chen[1,2,3]*

*Address all correspondence to: sien.chen@postgrad.manchester.ac.uk

1 Institute of Internet Industry, Tsinghua University, Beijing, China

2 Alliance Manchester Business School, University of Manchester, Manchester, UK

3 Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, China

References

[1] Blattberg RC, Deighton J. Manage marketing by the customer equity test. Harvard Business Review. 1996;(July-August):136-144

[2] Kumar V, Venkatesan R, Reinartz W. Knowing what to sell, when, and to whom. Harvard Business Review. 2006;**84**(3):131-137

[3] Bendapudi N, Leone R. Psychological implications of customer participation in co-production. Journal of Marketing. 2003;**67**(1):14-28

[4] DeSarbo W, Jedidi K, Sinha I. Customer value analysis in a heterogeneous market. Strategic Management Journal. 2001;**22**(9):845-857

[5] Porter M. Clusters and the new economics of competition. 76th ed. Boston: Harvard Business Review. 1998. pp. 77-90

[6] Gupta S, Lehmannand DR. Valuing customers. Journal of Marketing Research. 2004;**41**(1):7-18

[7] Dwyer FR. Customer lifetime valuation to support marketing decision making. Journal of Direct Marketing. 1997;**11**(4):6-13

[8] Berger PD, Nasr NI. Customer lifetime value: Marketing models and applications. Journal of Interactive Marketing. 1998;**12**(1):17-30

[9] Rust RT, Lemon KN, Zeithaml VN. Return on marketing: Using customer equity to focus marketing strategy. Journal of Marketing. 2004;**68**(1):109-127

[10] Blattberg EC, Malthouse FJ. Can we predict customer lifetime value? Journal of Interactive Marketing. 2005;**19**(1):2-16

[11] Rust RT, Zeithaml VA, Lemon KN. Driving Customer Equity: How Customer Lifetime Value Is Reshaping Corporate Strategy. New York: Free Press; 2000

[12] Blattberg RC, Getz G, Thomas JS. Customer Equity: Building and Managing Relationships As Valuable Assets. Boston: Harvard Business School Press; 2001

[13] Sheth J, Mittal B, Newman B. Consumer Behavior and Beyond. NY: Harcourt Brace; 1999

[14] Hennig-Thurau T, Hansen U. Relationship Marketing-Some Reflections on the State-of-the-Art of the Relational Concept. In: Hennig-Thurau T, Hansen U, editors. Relationship Marketing: Gaining Competitive Advantage Through Customer Satisfaction and Customer Retention. New York: Springer; 2000. pp. 3-27

[15] Fader PS, Hardie BGS, Lee KL. "Counting your customers" the easy way: An alternative to the Pareto/NBD model. Marketing Science. 2005;**24**(2):275-284

[16] Fader PS, Hardie BGS, Lee KL. RFM and CLV: Using CLV curves for customer base analysis. Journal of Marketing Research. 2005;**42**(4):415-430

[17] Schmittlein DC, Peterson RA. Customer base analysis: An industrial purchase process application. Marketing Science. 1994;**13**(1):41-67

[18] Reinartz WJ, Kumar V. The impact of customer relationship characteristics on profitable lifetime duration. Journal of Marketing. 2003;**67**(1):77-99

[19] Reinartz WJ, Kumar V. On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing. Journal of Marketing. 2000;**64**:17-35

[20] Bhattacharya CB. When customers are members: Customer retention in paid membership contexts. Journal of Academy of Marketing Science. 1998;**26**(1):31-44

[21] Bolton RN. A dynamic model of the duration of the customer's relationship with a continuous service provider: The role of satisfaction. Marketing Science. 1998;**17**(1):45-65

[22] Thomas JS. A methodology for linking customer acquisition to customer retention. Journal of Marketing Research. 2001;**38**(May):262-268

[23] Mulhern FJ. Customer profltability analysis: Measurement, concentrations, and research directions. Journal of Interactive Marketing. 1999;**13**(1):25-40

[24] Bell D, Deighton J, Reinartz J, Rust R, Swartz G. Seven barriers to customer equity management. Journal of Service Research. 2002;**5**(1):77-85

[25] Gupta S, Lehmann DR. Customers as assets. Journal of Interactive Marketing. 2003;**17**(1):9-24

[26] Jain D, Singh SS. Customer lifetime values research in marketing: A review and future directions. Journal of Interactive Marketing, 2002;**16**(2):34-46

[27] Cook RD, Weisberg S. Residuals and Influence in Regression. New York: Chapman and Hall; 1982

[28] Neter J, Kutner M, Nachtsheim C, Wasserman W. Applied Linear Statistical Models, 4th ed. Chicago: Irwin; 1996

[29] Buckinx W, Van den Poel D. Customer base analysis: Partial defection of behaviorally-loyal clients in a non-contractual fmcg retail setting. European Journal of Operational Research. 2005;**164**(1):252-268

[30] Hughes A. Strategic Database Marketing, 3rd ed. New York: McGraw-Hill; 2005

[31] Kumar V. CLV: A Path to Higher Profitability. Working Paper. Storrs: University of Connecticut; 2006

[32] Schmittlein DC, Morrison DG, Colombo R. Counting your customers: Who are they and what will they do next? Management Science. 1987;**33**(1):1-24

[33] Niraj R, Gupta M, Narasimhan C. Customer profltability in a supply chain. Journal of Marketing. 2001;**65**(3):1-16

[34] Venkatesan R, Kumar V. A customer lifetime value framework for customer selection and resource allocation strategy. Journal of Marketing. 2004;**68**(4):106-125

[35] Rossi PE, Allenby GM. Bayesian statistics and marketing. Marketing Science. 2003;**22**(3):304-328

[36] Venables W, Ripley B. Modern Applied Statistics with S-PLUS. New York: Springer; 1999

[37] Cui D, Curry D. Prediction in marketing using the support vector machine. Marketing Science. 2005;**24**(Fall):595-615

[38] Giuffrida G, Chu W, Hanssens D. Mining Classification Rules from Datasets with Large Number of Many-Valued Attributes. Computer Science. Vol. 1777. Berlin: Heidelberg; 2000. pp. 335-349

[39] Breiman L. Statistical modeling: The two cultures. Statistical Science. 2001;**16**(3):199-231

[40] Breiman L. Heuristics of instability and stabilization in model selection. Annals of Statistics. 1996;**24**(6):2350-2383

[41] Hastie T, Tibshirani RJ, Friedman JF. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer; 2001

[42] Glady N, Baesens B, Croux C. A modified Pareto/NBD approach for predicting customer lifetime value. Expert Systems with Applications. 2009;**36**(2):2062-2071

[43] 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 ACM, 2016. pp. 785-794

[44] Gupta S, Lehmann D, Stuart J. Valuing customers. Journal of Marketing Research. 2004;**41**(1):7-18

[45] Malthouse E, Blattberg R. Can we predict customer lifetime value?. Journal of Interactive Marketing. 2005;**19**(1):2-16