Customer Lifetime Value: An Ensemble Model Approach

Harminder Singh Channa

Abstract Customer lifetime value allows Banks and Financial Institutions to examine the worth of customers to the business, which provides important inputs to take informed marketing & retention decisions and better Customer Relationship Management. Traditional CLV approaches are primarily isolated at account level worthiness. Some Customer level CLV do take 360° view of the customer relationships but are more heuristic in nature or predicting the CLV based on historical CLV data using single model approach. In this paper, we have explored the existing solutions available to calculate the CLV and explained the rationale for not using with their respective limitations. The focus of the study was on retail banking sector, the proposal is to use whole gamut of existing marketing and risk predictive models for calculating the predicted CLV without taking the time value of money into consideration. It also discusses about the comparisons between the present and future CLV of the customer and how to check the overall health of the bank's business using calculated CLV.

Keywords CLTV · CLV · Customer lifetime value · Predicted customer value
Present customer value · CLTV limitations · Predictive model · Churn model
Survival model · Machine learning in banking · Customer potential
Time value · Present value · Customer relationship · Customer level CLTV

1 Introduction

CLV: Customer Lifetime Value in the most simplistic terms is the measure of the profitability of the customer during the lifetime. The limelight which CLV always receive is because all the banks are interested in focusing on the 'cream of the crop'

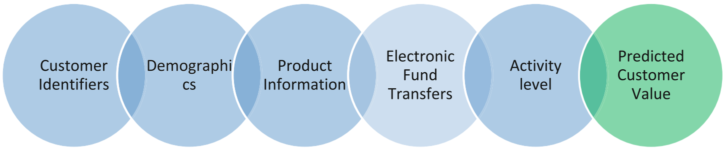
H. S. Channa (✉)
Analytics & Insights, Tata Consultancy Services, Bengaluru, Karnataka, India
e-mail: harminder.channa@tcs.com

V. E. Balas et al. (eds.), *Data Management, Analytics and Innovation*,
Advances in Intelligent Systems and Computing 808,
https://doi.org/10.1007/978-981-13-1402-5_27

customers to have sustaining mutually beneficial relationships with selected customers. Also, it exhaustively explains the relationship between the customer and the Bank.

This CLV soothsaying can vary from a subjective/heuristic approach to more sophisticated analytical techniques. There is no standard way of measuring CLV.

Bank and Data Sources
We were supporting an Indian bank[1] with strong regional presence and lot of data attributes were available to us to have a customized data lake. Besides model development for different bank departments, one of the prioritized goals was the calculate CLV for the customers.



In order to have our own CLV measure for the retail banking exhaustive enough to encompass all relationships and potentials of a customer, we started adopting one of the available approaches but ended up devising a novel way to calculate CLV.

Various approaches are in place and are published, focusing on multiple aspects in the retail banking industry. We have explained some of those and then boiled down to the solution we have presented.

Three standard approaches/techniques were explored that are used in industry:

(1) Churn model based approach
(2) Survival analysis based approach
(3) Model for CLV.

2 Literature Review

2.1 Approach 1: Churn Model Based Approach

CLV is viewed as the present value of the future cash flows associated with a customer. It is basically the sum of the incomes obtained from a customer over the lifetime after deducting all the associated costs by taking the present value of the money.

The basic formula for calculating CLV for customer, i at time, t for a finite time horizon T is:

[1]Due to proprietary bank, we are keeping the bank anonymous.

$$\mathrm{CLV}_{i,t} = \sum_{t=0}^{T} \frac{\mathrm{profit}_{i.t}}{(1+d)^t} \text{ or } \mathrm{CLV}_i = \sum_{t=1}^{T} \frac{\mathrm{Revenue}_{i.t}}{(1+d)^t} - \sum_{i=1}^{T} \frac{\mathrm{Cost}_{i.t}}{(1+d)^t} \tag{1}$$

Now calculating CLV based on T that is equal to the term of term deposits or advances would be naïve. That is to assume all the customers on the banks books would complete the full terms. For Saving/Current deposits selecting T becomes little tricky. A rule of thumb is to calculate it for next 3–5 years or any other horizon, which business feels would not change substantially.

But this approach with some fixed horizon is little vulnerable to customer attrition or churn. So an improved approach is to incorporate the churn prediction.

Here is the mathematical model for CLV measuring of this research:

$$\mathrm{CLV} = \sum_{t=1}^{n} \frac{P_t(S_t X M_t)}{d^t} - \sum_{i=1}^{n} \frac{(P_t X D_t) + R_t)}{d^t} \tag{2}$$

where

- P_t The probability of continuous interaction of customer with the bank; $P_t = 1 - \mathrm{C.R}$, and also C.R is churn probability.
- S_t The average amount of customer's accounts after subtracting by legal and liquidity saving rate; this amount of accounts inventory is the free deposits for retail banks.
- M_t The marginal profits for S_t.
- d_t Discount rate that is equal to: 1+ inflation rate.
- D_t This is the first group of costs that associated with the direct costs about the accounts.
- R_t This is the first group of costs that associated with the indirect costs. This group are including of costs such as: advertising and marketing costs, depreciation costs, administrative costs, other personnel costs, etc.
- N Number of periods.

2.2 Approach 2: Survival Analysis Based Approach

CLV based on survival function and time value of money.

Survival model is developed based on customer's past behavior and trends, to calculate the probability of a customer's survival for next "n" years. CLV is calculated based on historic and predictive Customer Lifetime Value for each customer.

$$\begin{aligned} \mathrm{CLV} &= \mathrm{CLV}_{\mathrm{History}} + \mathrm{CLV}_{\mathrm{Future}} \\ \mathrm{CLV}_{\mathrm{Future}} &= \mathrm{Survival}(t) * T * [\mathrm{Monthly\,Potential}] \end{aligned} \quad (3)$$

In another way, CLV represents the net present value from profits, from a single customer. It can be represented as

$$\mathrm{CLV}_{\mathrm{customer\_i}} = \sum_{i=1}^{N} (\mathrm{Revenue}_i \times \ \mathrm{Survivalrate}_i - \mathrm{Expenses}_i)/(1+r)^i \quad (4)$$

2.3 Approach 3: CLV Predictive Model

This approach is quite straightforward and is to develop a predictive model for Customer Lifetime value (CLV) using the most important variables from the banking industry.
This works like a standard model development approach with a dependent variable and lot of predictors. CLV is calculated for each customer in the historical data and used as a dependent variable. The dependent variable is the yearly profit obtained from the customer and is computed as taking the sum of the profits gained from each transaction, the assets and liabilities and the products/services used by the customers.

Various modeling techniques can be used to model CLV. This study for literature review did a comparison between least square estimation (LSE) and artificial neural network (ANN) in order to select the best performing forecasting tool to predict the potential CLV. Due to its higher performance; LSE based linear regression model is selected.

In addition to the common variables used in CLV prediction, monetary value and risk of certain bank services, as well as product/service ownership-related indicators, are also significant factors.

3 Proposed Work

As discussed above, we have first tried to adopt the approaches mentioned in the literature review but found that there was some room available to add value to the existing approaches. In this section, we first aimed to address the drawbacks of the approaches discussed through a simple but yet effective approach.

(1) Limitation with first two approaches:

Both the methods i.e. Churn & Survival based work well if CLV needs to be calculated at an account level. In order to take in every aspect of the customer relationship, the best we could do was to agglomerate individual CLVs at the customer level for different products. But it was missing out the entire customer potential. Customer value was also dependent on the future products/offers a customer can take up.

(2) Limitation with CLV predictive model:

Missing out on the customer potential and weighing so heavily on the historical predictors only, to predict the future CLV. It becomes a real challenge to squeeze in all the diverse customer information from all customer touch points/spheres into one model. This information ranging from demographics, experience, monetary and risk information to product/service ownership was so vast to be captured appropriately in one single model.

These limitations formed the basis of our research and drove us to work to encompass every aspect of the customer.

3.1 Predictive Models in Banks

From the literature review, one thing was certain that the predictive models, be it churn model or survival Cox proportional hazard model or Foreclosure model, do add a lot of value in CLV calculation/prediction.

Therefore, it became imperative for us to understand the current model repository a retail bank was equipped with and to assess whether we can leverage anything we have developed so far. The good part was that the retail banking industry has matured in the last few years and few banks managed to have evolved as a customer centric bank. This gave us a head start and following was how our banks model repository (Table 1) looked like for both Deposits & Advances[2] both marketing/risk models.

Table 1 Banks model repository

Marketing models	Risk models
Response & uplift model	Probability of Default & LGD Survival & Churn Model Payoff Model & Foreclosure

[2]Due to proprietary bank data, not all models were listed.

3.2 Usage of Predictive Models in Banks:

How banks use predictive models for marketing/Risk/Account management was an important area to look at.

We have observed, how much customer centric a bank can become, usage of models was mostly segregated or in isolation. What we mean here was various bank's departments mostly work in isolation.

- Marketing team's focus was always increasing the response rates and hence acquisitions. They have their own suite of models (like response, uplift, attribution) which they use without any knowledge what was happening to bank's NPAs
- Similarly Risk teams focus was to look into credit worthiness of the customers.

There was no problem working like this and that was a successful analytics model a bank can have but synergizing all departments would definitely yield better results.

Therefore, the available stock of models was good enough to have 360° view of the customers. Each model with individual probabilities predictions from all phases of customer lifecycle or customer potential and from all relationships are utilizing the entire data lake available.

3.3 Our Approach: Model Ensemble

Our approach was basically an ensemble approach where we were mainly "connecting the dots" to reveal a bigger picture where individual bank departments models were the dots.

> Our definition of CLV is the measure of the profitability of the customer during the lifetime from all the relationships and future potentials of a customer without taking into account the time value of the money by only incorporating available predictive models

$$\mathrm{CLV}_{\mathrm{customer},i} = \mathrm{CLV}_{\mathrm{customer}_i,\mathrm{Potential}} + \mathrm{CLV}_{\mathrm{Fcustomer}_i,\mathrm{future}}$$

where

$\mathrm{CLV}_{\mathrm{customer}_i,\mathrm{Potential}}$	It was the customer potential. It was revealed by the marketing or acquisition models predicting the future products a customer can take up.
$\mathrm{CLV}_{\mathrm{Fcustomer}_i,\mathrm{future}}$	It is CLV of existing relationships: Using all the available risk and retention models as used in account level CLV calculations.

Our definition prefers to have a shorter horizon of the customer relationship from 6 to 12 months as business conditions would not change drastically in this period and because of this there was no need to incorporate the time value of money.

3.4 Demonstration

From the Deposit & Advances books, this was the snapshot of two customers with the following relationships with the bank.

Cust_num: Customer Number
Month Year: Month and year when snapshot was taken
Deposit Amt: represents the term deposit account amount
Saving Curr bal: represents the saving account current balance
HLoan Appr Amt: represents the Home Loan account approved amount
HLoan bal Amt: represents the Home Loan balance amount
Loan Approved Amt: represents Personal loan account approved amount
Loan Bal: represents the Personal Loan balance amount

STEP 1 $\mathrm{CLV}_{,\mathrm{present}}$ (CLV at present)

Cust_num*	Month Year	Age	Region	Deposit Amt	Saving Curr bal	HLoan Appr Amt	HLoan bal Amt	Loan Approved Amt	Loan Bal	Present Customer Value
ABC	Aug15	35	Region Central	Rs0	Rs6,500	Rs15,000	Rs12,000	Rs5,000	Rs1,000	Rs15,000
HSC	Aug15	65	Region East	Rs0	Rs2,500	Rs0	Rs0	Rs2,000	Rs1,500	Rs2,650
.										
.										

*Masking the bank account number.

Based on the present portfolio of customers (ABC), $\mathrm{CLV}_{\mathrm{present}}$ was calculated i.e.

$$\mathrm{CLV}_{\mathrm{ABC,present}} = (\text{Deposit Amt}) + (\text{Saving Curr bal}) + 70\% * (\text{HLoan bal Amt}) + 10\% * (\text{Loan bal})$$

$$\mathrm{CLV}_{\mathrm{ABC,present}} = \mathrm{Rs.}15,000$$

STEP 2 Adding predicted probabilities per customer

With lot of predicted models in place, we had made these predictions about the customer

Cust_num	Age	Region	Deposit Amt	Saving Curr bal	HLoan Appr Amt	HLoan bal Amt	Loan Approved Amt	Loan Bal	Present Customer Value	p1	p2	p3	p4	Dec 1	Dec 2	Dec 3	LGD
ABC	35	Region Central	Rs0	Rs6,500	Rs15,000	Rs12,000	Rs5,000	Rs1,000	Rs15,000	0.4	0.0055	0.06	0.15	3	10	2	0.15
HSC	65	Region East	Rs0	Rs2,500	Rs0	Rs0	Rs2,000	Rs1,500	Rs2,650	0.56	0.008	NA	0	2	9	NA	0.0001
.																	
.																	

*p*1 Represents the probability of existing customers buying a Term Deposit

*p*2 Represents the Churn probability of existing Saving customers within next 12 months

*p*3 Represents the payoff probability of existing Home Loan customers to foreclose

*p*4 Represents the predicted Loss Given Default of existing Personal Loans within next 6 months

Dec1 Represents the predicted decile of existing customers buying a Term Deposit

Dec2 Represents the predicted Churn decile of existing Saving customers within next 12 months

Dec3 Represents the predicted payoff decile of existing Home Loan customers to foreclose

LGD Represents the predicted Loss Given Default of existing Personal Loans within next 6 months.

STEP 3 Based on predicted probabilities/deciles per customer, predicted amounts are calculated, which finally converted into CLV

Cust_num	Age	Region	Present Customer Value	p1	p2	p3	p4	p5	Dec1	Dec2	Dec3	LGD	RsVal1	RsVal2	RsVal3	RsVal4	Predicted Customer Value
ABC	35	Region Central	Rs15,000	0.4	0.0055	0.06	0.15	0.002	3	10	2	0.15	Rs750	Rs6,500	(Rs8,400)	(Rs150)	(Rs1,300)
HSC	65	Region East	Rs2,650	0.56	0.008	NA	0	0.0005	2	9	NA	0.0001	Rs750	Rs2,500	Rs0	(Rs0)	Rs3,250

RsVal1 Represents the customer's potential to buy Term deposit. We have used average term deposit = Rs. 750

RsVal2 Represents the customer's potential loss as churning from Saving account, which was equivalent to the current saving balance. In the example, customer was not churning therefore, current balance i.e. Rs. 6,500

RsVal3 Represents the customer's potential loss on Home loan outstanding balance as prepaying the loan. Calculated as (−1)** × 70% of Rs. 12000 = −Rs. 8,400

RsVal4	Represents the customer's potential loss on Personal loan outstanding balance as customer will default. Calculated as (−1)** × LGD of Rs. 1,000 = − Rs. 150
Predicted customer value	It was the sum of the potential gain/loss per customer.

Calculated as $CLV_{ABC,predicted}$ = Rs. 750 + Rs. 6,500 + (−Rs. 8,400) + (−Rs. 150) = −Rs. 1,300.

**(−1) to subtract the loss

3.5 *Comparison Between Present and Future CLV*

Therefore, Customer who seemed very profitable as on Aug15 with $CLV_{ABC,present}$ = Rs. 15,000 has the predicted $CLV_{ABC,predicted}$ of – Rs. 1,300 as depicted in Fig. 1.

In order to pin point the concern areas, we calculated the percentage change in each of the customer products, which was represented in the Fig. 2.

Predicted CLV of the Bank: In the demonstration above, we had calculated the predicted CLV of one customer. When we replicated it across the Deposits and Advances book, we got astonishing results. The healthy and growing looking book based on the present value has shown inverted picture as shown in Fig. 3.

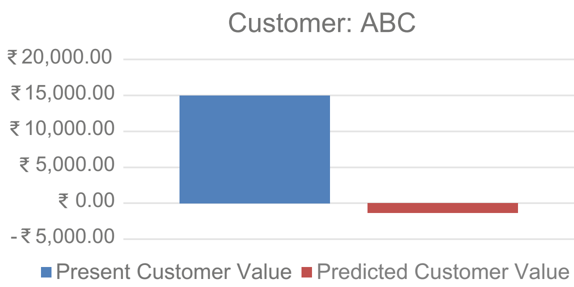


Fig. 1 Present versus predicted Customer value

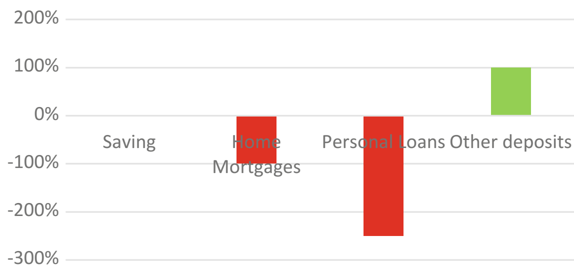


Fig. 2 Portfolio wise: percent change

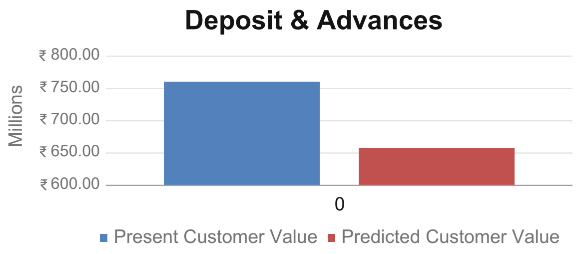


Fig. 3 Present versus predicted Customer value of deposits & advances

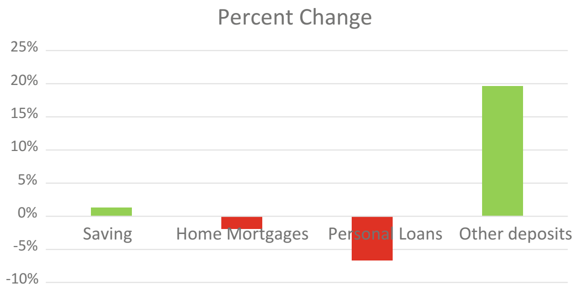


Fig. 4 Portfolio wise across bank: percent change

In order to pin point the concern areas, we calculated the percentage change in each of the customer products for the bank's Deposit and Advances, which was represented in the Fig. 4.

This gave a very important insights:

(1) Advances were going to give losses in the near future
(2) Deposits were expecting growth of 20%.

4 Conclusion

The proposed approach was very easy to adopt and implement and captures all the important parameters available in the retail banking space. It not only answered the problem CLV prediction questions but also gave important insights about how marketing and retention strategies at a customer level. The highlight was to get enlightened about the overall health of the business. It very easily explained the focus area to start prioritizing with to understand from which segment the maximum impact was coming.

References

1. Berger, P., & Nasr, N. (1998). Customer lifetime value: marketing models and applications. *Journal of Interactive Marketing, 12*(1), 49–61.
2. Kahreh, M. S., et al. (2014). Analyzing the applications of customer lifetime value (CLV) based on benefit segmentation for the banking sector. *Procedia—Social and Behavioral Sciences, 109,* 590–594.
3. Ramanathan, D. (2014–2015). How to become a CLV Aligned Organization: Fractal analytics.
4. Ekinci, Y., Uray, N., & Ülengin, F. (2014). Lifetime value in the bank industry: Sixth montreal industrial problem solving workshop August 17–21, 2015. *European Journal of Marketing; Bradford, 48*(3/4), 761–784.