Intact Workshop Report

Tommy Mastromonaco, Zehai Wen, Kylian Ajavon, Arsene-Brice Zotsa-Ngoufack

August 2023

1 Introduction

The insurance industry is a dynamic sector that operates in a high-stakes environment. It involves the assessment and management of numerous variables, ranging from risk evaluations to understanding customer behavior. In this context, the ability to gain actionable insights is crucial.

The **Customer Lifetime Value** (or CLV) refers to the total expected profit a company expects from a client throughout their entire relationship. It is an important tool for navigating the complexities of the insurance industry. At its core, the CLV serves as a unified metric that encompasses several variables, allowing insurers to better understand, predict, and manage the information they have about their clients. As a result, the CLV facilitates the decision-making process and allows insurers to make more informed decisions about their clients.

It can be modeled as follows:

$$CLV(a) = \mathbb{E}\left[\sum_{t=1}^{T} \gamma^t Profit(S_t) \mid S_0 = a\right], \tag{1}$$

where γ is a discounting factor to account for time-value of money, and $Profit(S_t)$ is a function that gives the expected profit from a client given their state S_t.

2 Method

First and foremost, we need to model $\{S_t\}$ in order to compute the CLV. A simple approach is to view $\{S_t\}$ as a sequence of random variables and assume that $\{S_t\}$ satisfies the Markov property, which means that

$$\mathbb{P}(S_{t+1} = s \mid S_t, S_{t-1}, \ldots, S_0) = \mathbb{P}(S_{t+1} = s \mid S_t). \tag{2}$$

With this assumption, the CLV can be computed using a method described in [1] with the following three steps:

1. Fit a regression tree on the data to identify groups (i.e. the states of the Markov chain) using the profit as a target variable;

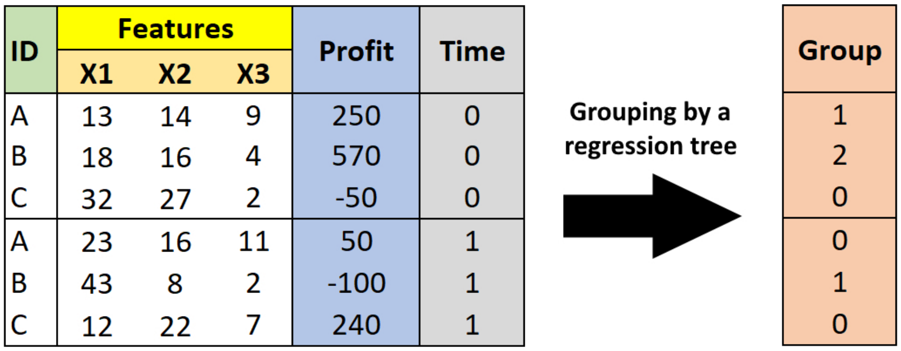
2. Estimate the transition probabilities between each group/state;

3. Compute the CLV by using the Monte Carlo method.

Let's dive into the method and illustrate it with a basic example, which involves a portfolio of three customers labeled as **A**,**B** and **C**, and observed on $t = 0, 1, 2$. Each observation consists in feature variables labeled as **X1**, **X2** and **X3**, and a **Profit** variable.

Step 1 Combine the data from all time periods into one dataset (assuming that customer characteristics are time-independent) and use this dataset to fit a regression tree.

After merging all the data, the regression tree divides the space of features into classes or groups based on the profit variable. Once the classes are formed (groups 0, 1 and 2 in our example), a new feature named **Group** is created by determining which group each observation belongs to, as shown in figure 1 (only $t = 0, 1$ are shown to lighten the illustration). Note that this variable is ordered, since the tree associates a mean profit with each group.



ID	Features			Profit	Time
	X1	X2	X3		
A	13	14	9	250	0
B	18	16	4	570	0
C	32	27	2	-50	0
A	23	16	11	50	1
B	43	8	2	-100	1
C	12	22	7	240	1

Grouping by a regression tree

Group
1
2
0
0
1
0

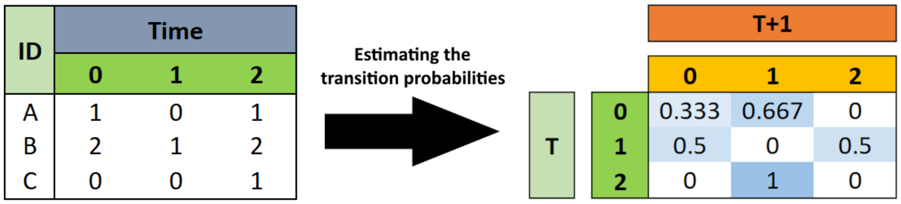
Figure 1: Grouping using a regression tree

Step 2 Build a transition matrix from empirical transition probabilities, assuming the Markov chain is homogeneous:

$$p_{ij} := \mathbb{P}(S_{t+1} = j \mid S_t = i) = \mathbb{P}(S_t = j \mid S_{t-1} = i) \quad \forall t. \tag{3}$$

The assumption of time-homogeneity is very helpful since it allows to compute a single transition matrix instead of having to compute a different matrix for each time interval. Thus, the empirical transition probabilities are

$$\hat{p}_{ij} = \frac{\#\text{ of transitions from } i \text{ to } j}{\#\text{ of transitions from } i}. \tag{4}$$



ID	Time		
	0	1	2
A	1	0	1
B	2	1	2
C	0	0	1

		T+1		
		0	1	2
T	0	0.333	0.667	0
	1	0.5	0	0.5
	2	0	1	0

Figure 2: Estimating the transition matrix

Figure 2 illustrates the computation of the transition probabilities in our example. On the left are shown the transitions between groups and on the right the resulting transition matrix. If we consider, let's say, $\hat{p}_{01}$, we see on the left that there are three transitions from group 0, two of which lead to group 1. Hence $\hat{p}_{01} = 2/3$.

Step 3 Compute the CLV by simulating Markov chains with the Monte Carlo method.

In order to compute $CLV(a)$, we simulate multiple paths of $\{S_t\}$ starting from $S_0 = a$ by using the empirical transition matrix. If for the i^{th} path out of N paths the computed CLV is $\widehat{CLV}_i(a)$, then by the Monte Carlo method, the estimated CLV is

$$\widehat{CLV}(a) = \frac{1}{N}\sum_{i=1}^{N}\widehat{CLV}_i(a). \tag{5}$$

3 Results

We implement the above methods to synthetic data. The synthetic data has less features compared to the real data. A generalized linear model is given to compute the profit of each client. Only the features necessary for the algorithm is present. Lots of other features, especially the categorical ones, are omitted in the synthetic data. Each synthetic feature is generated uniformly from the bounds of the corresponding features of the real data. We expect that:

1. Our implementation of decision tree gives a reasonable number of states.
2. The CLV values of all states should be bounded reasonably.
3. The CLV values of all states should be random and centred around zero.

Such is indeed what we observe:

Attempts are made to the real data after the success on the synthetic data. However, due to programming difficulties, cleaning the real data alone took too much time and we were not able to get results on the real data due to time constraint.

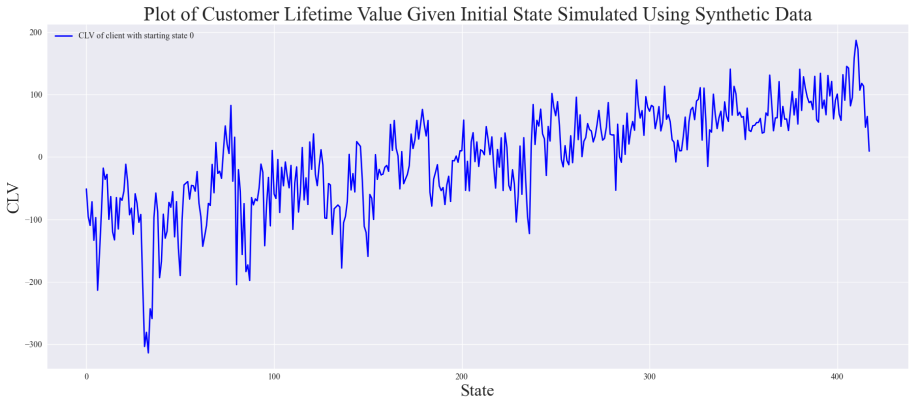


Figure 3: Simulation of the CLV

4 Discussion

In this section, we look at other approaches to calculating CLV.

In the literature, CLV has been extensively studied in both academic research and in companies for marketing purposes. One of the most famous references is that of Gupta [2], who presents methods for modeling CLV. There are several of these, including the markov chain approach, which consists first of all in determining the various possible customer states using various algorithms, such as the regression tree used here, or the Kmeans method, and then calculating the CLV using the formula (1) , we refer to [3,4] for more details. On the other hand, we also have the pareto/NBD model, a famous model of the CLV. This model assumes that recency and frequency are independent of monetary value, and is divided into two sub-models: one for the expected number of transactions, the other for the expected average order value. Multiplying the results gives CLV. We refer to [3] to the assumptions for the transaction pareto/NBD submodel.

References

[1] Haenlein, M., Kaplan, A. & Beeser, A. (2007). *A Model to Determine Customer Lifetime Value in a Retail Banking Context.* European Management Journal (Vol. 25).

[2] Gupta Sunil, Hanssens Dominique, Hardie Bruce, Kahn Wiliam, Kumar V., Lin Nathaniel, Ravishanker Nalini and Sriram, S. *Modeling customer lifetime value.* Journal of service research, 9.2 (2006): 139-155.

[3] Jasek, P., Vrana, L., Sperkova, L., Smutny, Z., & Kobulsky, M. (2018). *Modeling and application of customer lifetime value in online retail.* Informatics (Vol. 5, No. 1).

[4] Jablecka, M. (2020). *Modelling CLV in the Insurance Industry Using Deep Learning Methods.* Master's Thesis at KTH Royal Institute of Technology.