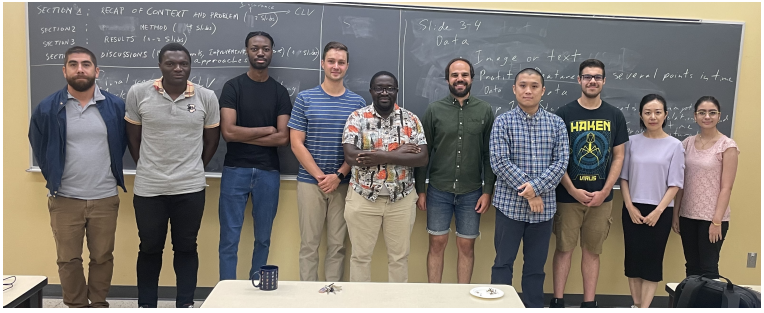
Estimation of the Customer Life Value



August 25, 2023

Insurance context

- **Providing Insights in a Complex Industry:**
  - Insurance operations involve numerous variables, from risk assessment to customer behavior.
  - **Customer Lifetime Value** (or CLV) offers a comprehensive metric encompassing these factors.
- **Efficient Decision-Making:**
  - CLV consolidates diverse information, streamlining decision processes.
  - Enables optimized resource allocation, customer engagement, and tailored product offerings.

Customer Life Value (CLV)

- CLV represents the total expected profit a company expects from a client throughout their entire relationship.
- Used in multiple industries in order to evaluate the financial value of a customer and better tailor the approach of the company towards customers (pricing, marketing, etc.)
- Mathematically, we can define CLV as

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

where:

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

The model

- Problem: how to model S_t?
- Natural to think of $\{S_t\}$ as a sequence of r.v.
- We assume the Markov property for simplification:

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

We used a method from Haenlein et al. (2007) that involves 3 steps:

1. Fit a regression tree on the data to identify groups (i.e. the states of the Markov chain) with the profit as a target variable;
2. Estimate the transition probabilities between each group/state;
3. Compute the CLV by Monte Carlo.

Details

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

Group
1
2
2
0
1
0

- **Step 1**: Combine data from all time steps into a single dataset (we assume time independency) and fit a regression tree;
- Result : that creates a new feature **Group** (there is a sense of order by profit). We can "forget" the other features from now on.

Details (continued)

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

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

- **Step 2**: Build the transition matrix with empirical transition probabilities (assuming time homogeneity);
- **Step 3**: Compute the CLV by simulating Markov chains (Monte Carlo method).

Other approaches

Extended Pareto/NBD Model

- Combines **Pareto/NBD** and **Gamma-Gamma** models.
- **Pareto/NBD Model**:
  - Uses two main components: Pareto and Negative Binomial Distribution.
  - Estimates parameters that describe customer behavior, such as transaction rate and the expected number of future transactions.
- **Gamma-Gamma Model**:
  - Assumes that customer transaction values follow a gamma distribution.
  - Estimates parameters for average and variability in transaction values.
- Gives **CLV predictions** by multiplying results from both submodels.
- Challenges Faced:
  - Attempted implementation, but faced technical hurdles.
  - The dataset provided for the workshop was not suited for implementation (missing variables).

References

- Haenlein, Michael & Kaplan, Andreas & Beeser, Anemone. (2007). *A Model to Determine Customer Lifetime Value in a Retail Banking Context.* European Management Journal. 25. 221-234. 10.1016/j.emj 2007.01.004.