RETINA

Retina’s Approach to Customer Lifetime Value

RETINA.AI

Contents

Who is Retina?

We are your CLV experts. Retina is the customer intelligence partner that provides accurate customer lifetime value metrics early in the customer journey. Optimize marketing budgets in real-time, drive more predictable repeat revenue, and elevate brand equity with CLV.

Retina maximizes value by helping you attract more high-value customers, avoid unprofitable customers, and boost lifetime value for those in the middle. We calculate CLV early in the customer journey to inform your strategies and predict your future valuation.

In this whitepaper, we will cover the following **concepts**.

1. RFM & BTYD Models
2. How to Estimate and Simulate Customer Behavior
3. Business Applications of CLV

Introduction

When it comes to mid- and long-term decision making, CLV is a reliable customer metric—more so than revenue—because it normalizes customer spending over time. Otherwise, recently acquired customers collectively appear to spend less than older customers.

Yet CLV data science thwarts many technical teams. It's challenging to predict individual behavior in a changing marketplace, for customers with little or no history, and to do so at scale.

In this whitepaper, Retina introduces a newly launched CLV framework designed to handle all these requirements and beyond. We build upon classical RFM techniques while leveraging cutting edge semi-supervised learning algorithms.

What's more—our models are interpretable and power many CLV-adjacent products designed to add color, context and precision to your CLV-driven decisions.

Retina's CLV is early, accurate, scalable and flexible. Use it to:

- stop paying for bad customers,
- project revenue and inventory demand into the future,
- frame business development decisions using customer-level economics, and
- track long-term goals by adding CLV metrics to company KPIs.

CLV Basics

RFM models. Let's start on common ground with a quick review of RFM models and cohort-based CLV. This is an intuitive and widely-known approach. The idea is to use customer-level

- recency (R),
- frequency (F), and
- monetization (M)

in order to

- (i) predict customer churn, and
- (ii) simulate the future behavior.

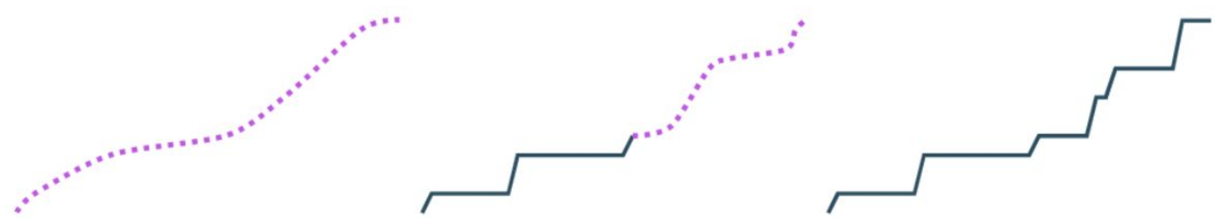
Models that build on this type of behavior are known by the acronym RFM.

When we make the following assumptions:

- customers are homogenous (behave identically to one another),
- churn is memoryless, and
- marketplaces do not change, then

cohort CLV (we refer to this as LTV) can be calculated on the back of an envelope.

CLV over time. What's missing is individual-level resolution and the timing of future behavior. This level of precision is critical for personalized decision making. For the rest of this whitepaper, imagine every customer—including prospective customers targeted for acquisition—has their CLV mapped out over time as seen in Figure 1.



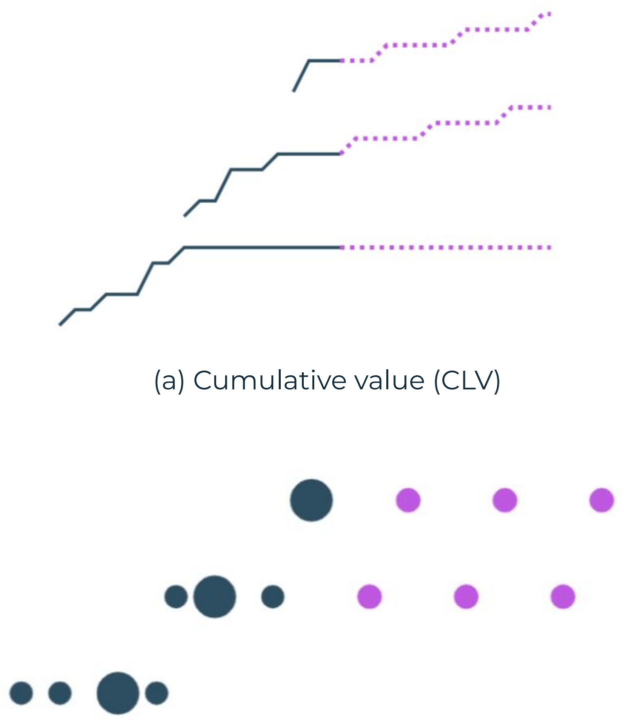
(a) prospect customer (b) active customer © churned customer

Figure 1: CLV through a lifecycle where — is known and •••• is estimated.

CLV for churned customers (who no longer add value) is known from data records. On the other hand, CLV for prospect customers (who might be targeted or anti-targeted for acquisition) is not known and must be estimated from other customers. Meanwhile CLV for currently active customers lies somewhere in between: value from the past is known while value in the future must be predicted.

Here’s a simple way to extend our back-of-the-envelope cohort LTV to customer-level CLV over time. Calculate averages for revenue-per-order and time-between-orders (i.e., intra-transaction time or ITT). Customers whose recency exceeds average ITT are labeled as churned. Compute cohort churn rate using the total number of transactions and the number of churned customers. From our memoryless-churn assumption, the average orders-until-churn for every active customer will be the fractional inverse of this churn rate.

Use these estimates to predict timing, size and duration of future orders for active customers, using recency to offset time, as seen in Figure 2. Subfigure (2a) shows the lifetime value (a cumulative sum) as time progresses for three separate customers. Meanwhile subfigure (2b) shows the incremental value (the orders themselves) for those same customers to match the cumulative totals



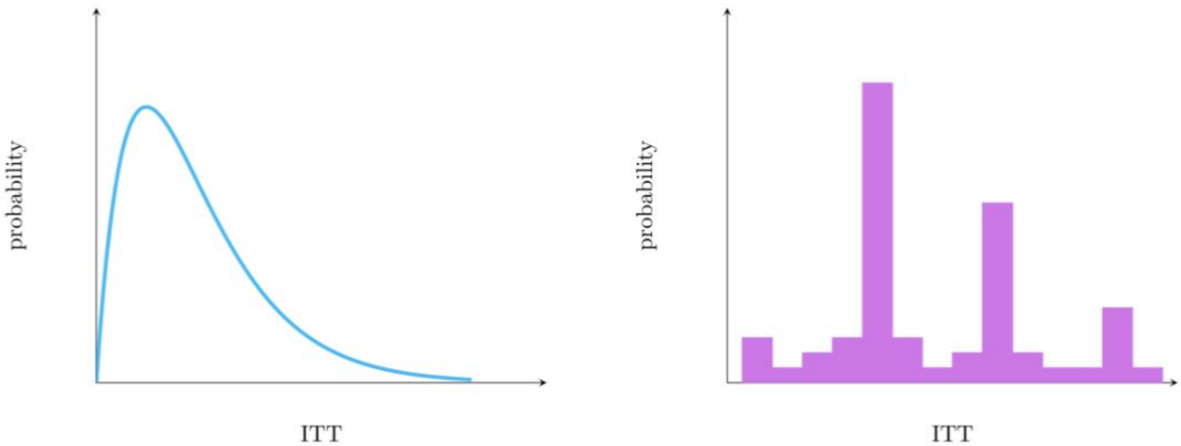
(a) Cumulative value (CLV)

(b) Incremental value

Figure 2: CLV using RFM for three customers, one of whom is churned.

BTYD models. These ideas have taken root and flourished in the Marketing Science community—inspired by salient Bayesian Statistics influences—into a family of sophisticated probabilistic models known as Buy Til You Die (BTYD) models. BTYD models are loaded with parametric-family assumptions, a hallmark of Bayesian paradigms.

But customers with regular buying patterns (say, a subscription service) accompanied by skips and one-time purchases will have ITT distributions that appear bi-modal or tri-modal, as seen in Figure (3b). This violates the prevalent probabilistic assumptions in BTYD models: that ITTs be drawn from a family of distributions—such as negative binomial or gamma—as seen in Figure (3a).



(a) Gamma (continuous) distribution

(b) Nonparametric (discrete) distribution

Figure 3: Customer ITTs may look like (b) while BTYD assumes they're drawn from (a).

Flavors and variations. To conclude our review of CLV basics, let's recognize the varying characteristics of CLV: flavors arise from variations in (1) naming, (2) value, and (3) scope. We've seen many different names—LTV, LTR, CLTV, PLTV, etc.—all reference the same set of ideas: customer behavior, prediction, and analytics over time. We've seen "value" defined differently: it may refer to revenue, profit, product interest, viewing time, referrals, or something else. CLV may refer to customer value that's been aggregated over individuals, time, both, or neither. Finally, predictions are often made for years at a time, but this level of granularity may vary as well.

In this whitepaper, we'll tell a simple yet common story about a metric called CLV: events are transactions, value is un-discounted revenue, and customer-level predictions project each week into the future.

Retina's Approach

At a high level, we use a new two-step process from RFM models:

(i) estimate customer-level patterns of behavior, in order to
(ii) simulate behavior into the future.

For each step, we'll discuss our design choices and how they map to requirements for CLV that is early (in a customer lifecycle), accurate, scalable, and flexible (easily convertible to any flavor).

Enhance how orders are simulated. What distinguishes BTYD from RFM models is the treatment of ITT. The customer frequency (F) statistic from RFM is a single outcome, while its intra-transaction time (ITT) counterpart from BTYD is a random variable over many outcomes. Traditional statistical models assume ITT PMFs are drawn from a family of parametric distributions, in order to use well-established tools. This exact decision is what unites most BTYD models throughout decades of research.

By contrast, we will not make this assumption because we see ITTs that violate this in practice. (Refer to Figure 2.) Our intuition is that flexible, nonparametric ITT distributions (i.e., probability mass functions or PMFs) will allow us to model reality more closely than BTYD. What's more, we'll use a new, semi-supervised learning model to handle this relaxation that offers computational advantages as well.

Use semi-supervised learning. Semi-supervised learning is a type of machine learning where

(i) there are no target signals, because the goal is pattern extraction, and/or
(ii) the target signals cannot be trusted at face value and must be processed.

By contrast, supervised learning and neural networks can be sometimes used to predict CLV in settings where the future horizon is short, such as gaming. While interesting, we focus on CLV models that are capable of arbitrarily long-term prediction horizons.

Our learning framework can impute (i.e., estimate using context clues) unknown ITT PMFs for hypothetical customers, and denoise the ITT PMFs for customers with too little order history. What's more, there's usually additional customer data from which to learn. We use Generalized Low Rank Models (GLRMs) that learn from all available customer data.

Learn from all customer data. Right away, we organize all data into a single Customer Table, where rows correspond to individual customers and columns to features. Table 1 illustrates what a Customer Table may look like.

Customer ID	Snack	Wine	AOV	churn %	ITT PMF
u12475016	nuts	red	$17.65	36%	
u15612916		both	$9.42	13%	
u51294191	coffee	white		28%	
u52318112			$31.12	71%	
u256185125	coffee	red		28%	
u335127169	cheese		$22.15	39%	
⋮	⋮	⋮	⋮	⋮	⋮

Table 1: Customer Table with **RFM** features and empirical PMFs (—).

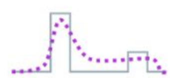
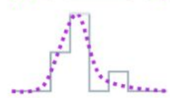
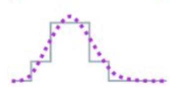
Customer ID	Snack	Wine	AOV	churn %	ITT PMF
u12475016	nuts	red	$17.65	36%	
u15612916	cheese	both	$9.42	13%	
u51294191	coffee	white	$31.25	28%	
u52318112	nuts	red	$31.12	71%	
u256185125	coffee	red	$31.25	28%	
u335127169	cheese	both	$22.15	39%	
⋮	⋮	⋮	⋮	⋮	⋮

Table 2: Customer Table with **imputed** valued and denoised PMFs(**...**).

Customers will appear in some data sets and not others. During the consolidation process, this surfaces in the form of missing information, and that's no problem. In fact, any data that's missing or unknown may be imputed using GLRM as part of semi-supervised learning. Retina's augmented results are shown in Table 2.

Of interest is the question of how to denoise and impute ITT PMFs. In fact, our choice to use a GLRM is rooted in its ability to specify loss functions that measure a dissimilarity of fit. For instance, we use Wasserstein (i.e., Earth Mover's) distance to handle ITT PMFs.

Solve for future outcomes in expectation. Now that the Customer Table is denoised and imputed, what remains is to simulate behavioral outcomes. Guessing a single customer's incremental value is easy:

- calculate the probability that another transaction occurs,
- draw a random ITT for time to next order, and
- assume the event will have average order value.

Continue this procedure until enough time has elapsed, and then aggregate over the simulated events.

Throughout the simulation, random events are assumed to be independent and identically distributed (IID). As the number of random simulations grows, the more their average starts to resemble a reliable prediction.

Using basic probabilistic principles, we've found an analytical solution that calculates expected customer value as a function of time—even when ITTs are nonparametric distributions. The key lies in our assumption that there is a maximum number of events that may happen in some fixed period of time.

Structure into a Customer Learning Framework. Let's reflect. We learn customer RFM signals from a GLRM that imputes and denoises a Customer Table. Fortunately the system inherits many of the useful properties that make GLRMs so powerful:

- the models are interpretable,
- there's no limit to the prediction horizon,
- solution methods are engineered to be fast and parallelizable, and
- also support small batches and online learning.

Figure 4: Retina's CLV framework follows three stages, each with its own validation phase.

CLV in Business

At Retina, we offer data enrichment via CLV and decision support for business solutions that leverage CLV. Here are some examples of CLV in action, with more details available in our case studies.

Growth marketing. With today's targeted advertising tools, overspending on bad customers is an avoidable mistake. The commonly used Return On Ad Spend (ROAS) metric uses spend-at-conversion as a proxy for customer value. Paired with counterpart Customer Acquisition Cost (CAC), ROAS serves the immediate need to measure campaign value, but it will be misleading in the long-run.

Rather, target high-CLV customers with lookalike audiences. Facebook and Google make it easy to upload customers with an associated value (e.g., CLV) to serve as seeds for lookalike audiences.

As confidence in CLV grows, adjust budgets to optimize for CLV, CAC, risk, acquisitions, and other factors. You can optimize budgets at the customer-level (a method called value-based bidding) or the campaign-level (CLV-optimized budget allocation).

Fiscal and inventory projections. Use CLV to manage inventory and supply chains: treat "value" as product demand, aggregate over customers, and peer into the near future. Alternatively, stick with revenue and catch underperforming strategies weeks earlier. Just remember that CLV in this setting only accounts for current customers, and additional business due to growth would need to be modeled separately.

We've seen this work well in settings where product offerings are limited, follow a subscription model, and offer skips and one-off purchases. We suspect the semi-periodic subscription service enables precise week-level forecasting while individual variations in skips and buys are learned then reflected in the CLV.

Customer success and fulfillment. A common assumption is that personalized attention leads to healthy relationships with the business. Yet promo-abusers and fastidious consumers may cost your business in the long run if appeasements and business costs outweigh revenue. So personalize appeasements and cap budgets to reflect what CLV remains from each customer.

Also it's common to have periods of overwhelming traffic, where tickets need to be prioritized and others dropped. The simplest way to handle ticket prioritization is to order tickets by the future (remaining) associated CLV, then start with the largest and work downwards.

During periods of supply chain and demand disruption (for example, COVID-19), there may even be a need to allocate fulfillment orders to customers with the highest CLV, then swap, delay, or reject orders from customers with the lowest CLV.

Unit economics and analytics. A burgeoning paradigm questions the use of order-level unit economics in favor of customer-level unit economics. This preference grows in proportion to the ratio of CAC to margin: as marketing costs begin to outweigh profit margins, costly customers grow more untenable. More broadly, CLV analytics reveal customer-level aspects of the business.

There are many simple and intuitive diagnostic reports for showcasing customer health. These include (but are not limited to):

- customer retention charts,
- sales concentration charts,
- historical revenue with predicted trajectories,
- customer segments by CLV, and
- payback period monitoring.

For further reading, check out our sales concentration whitepaper or inquire about our Quality of Customers report (an analog to a Quality of Earnings report).

Conclusion

In this whitepaper we started with the basics of customer lifetime value by looking at RFM and BTYD models, and then reviewed the various flavors of CLV. Next, we shared a new process to model CLV that estimates customer-level patterns in order to simulate behavior into the future. Finally, we discussed business applications of CLV, including growth marketing, fiscal and inventory projections, customer success and fulfillment, and unit economics and analysis.

CLV is a reliable customer metric for decision making because it normalizes customer spending over time. However, it's a challenge to predict accurately CLV, particularly early in a customer's journey. At Retina, we've built a CLV framework to do just that.

References

M. Udell, C. Horn, R. Zadeh, and S. Boyd. Generalized low rank models. *Foundations and Trends in Machine Learning,* 9(1):1-118, 2016.

M. Platzer and T. Reutterer. Ticking away the moments: TIming regularity helps to better predict customer activity. *Marketing Science*, 35(5):779-799, 2016.

D.G. Morrison Schmittlein, D.C. and R. Colombo. Counting your customers: Who are they and what will they do next? *Management Science*, 33(1):1-24, 1987.

A. Guitart A. Fernandez del Rio Chen, P. and A. Perianez. Customer lifetime value in video games using deep learning and parametric models. Pages 2134-2140, 12 2018.

C. Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

At Retina, we partner with our clients to predict customer lifetime value early in the customer journey. Optimize campaign budgets in real-time, drive repeat revenue, and boost customer and business health.

A few of our satisfied customers include:

If you're interested in partnering with our team of experts to bring the power of CLV to your business, schedule a demo to see what Retina can do for you.

RETINA.AI