Customer Lifetime Value Optimization and Retention Strategy: A Survival Analysis Approach to Telecom Churn Management

Julian Oppedisano

Hannah Wang

Jimmy Chu

MRKT 671: Advanced Marketing Analytics
Summer 2025

Introduction

Problem Description

The telecommunications industry faces unprecedented challenges in customer retention, with annual revenue loss due to churn reaching $34.4 million in our dataset analysis. Our data reveals a churn rate of 49.56%, significantly above the industry average of 15-25%, while average monthly revenue per customer stands at $57.91. With customer acquisition costs 5-25 times higher than retention costs, the economic imperative for effective churn management is clear. Traditional approaches treat churn as a uniform binary outcome, applying generic retention strategies that fail to account for customer heterogeneity and the time-dependent nature of customer defection.

This study addresses the fundamental question: How can telecom providers optimize customer lifetime value through integrated survival analysis, machine learning, and causal pricing strategies that account for customer heterogeneity and temporal churn dynamics? Our objective is to develop a comprehensive framework that transforms reactive churn management into proactive value optimization by identifying high-risk, high-value customer segments and implementing targeted interventions that maximize long-term profitability.

Plan of Action

Our scientific approach employs a three-stage methodology combining machine learning, survival modeling, and causal inference.

Stage 1: Ensemble Machine Learning

We implement ensemble machine learning using XGBoost, LightGBM, and CatBoost primarily for feature importance analysis and dimensionality reduction. Using SHAP values for interpretable feature ranking (Figure 1), we identify key predictors while addressing multicollinearity issues inherent in the 203-variable dataset. This stage serves as sophisticated feature selection rather than primary prediction, ensuring our subsequent survival models focus on the most informative variables.

Stage 2: Survival Analysis

- Cox Proportional Hazards Model:
  $h(t|X) = h_0(t) \times \exp(\beta_1 \mathrm{age} + \beta_2 \mathrm{calls} + \beta_3 \mathrm{equipment\_age} + \ldots)$

We conduct survival analysis using Cox Proportional Hazards modeling as our primary predictive framework, properly handling censored observations where 49.56% of customers haven't churned. Key predictors include customer age ($\exp(\beta) = 0.870$, $p < 5\times10^{-51}$), total calls ($\exp(\beta) = 0.726$, $p < 1\times10^{-302}$), and equipment age. Our Kaplan-Meier estimation (Figure 2) reveals paradoxical findings where high-revenue customers churn faster than low-revenue customers, challenging conventional wisdom about customer value and loyalty relationships.

Stage 3: Causal Pricing Model

- **Stage 1 - Choice Model**:
  $P(\text{customer } i \text{ chooses plan } j) = \exp(V_{i}\square) / \Sigma\square \exp(V_{i}\square)$

- **Stage 2 - Survival Integration**:
  $h(t|X_i) = h(t) \times \exp(\beta X_i + \gamma E[price_i])$

Third, we implement a two-stage causal pricing model combining conditional logit choice modeling with survival integration. Stage 1 estimates customer plan choice probabilities, while Stage 2 integrates expected prices into survival hazard functions. Our empirical estimation focuses on: (1) time-varying churn hazard rates as functions of observable characteristics, (2) customer lifetime value incorporating survival probabilities over 60-month horizons, and (3) optimal pricing parameters through counterfactual analysis that maximizes revenue while maintaining acceptable churn levels.

Empirical Section

Data Explanation

Our analysis utilizes a comprehensive telecom dataset containing 100,000 customers with 203 variables spanning demographics, service usage, billing, and churn indicators. Key variables include months, rev_Mean (average monthly revenue), mou_Mean (monthly minutes), totcalls, custcare_Mean (customer care interactions), age1, and eqpdays (equipment age). Data collection occurred across July, September, November 2001, and January 2002, focusing on mature customers with ≥6 months tenure.

Critical data characteristics inform our modeling strategy. The balanced churn outcome (49.56% churn rate) eliminates class imbalance concerns. Equipment age demonstrates strong predictive power (4.8% importance in ensemble models), while customer care interactions paradoxically correlate with higher churn risk rather than satisfaction. Most significantly, we observe a counterintuitive positive correlation between usage intensity and churn probability, indicating that high-engagement customers may actually represent higher risk segments requiring targeted retention efforts.

Key Estimation Results

Our ensemble machine learning model achieves superior performance with 80.73% recall for identifying churning customers. Feature importance analysis confirms equipment age as the primary predictor (highest SHAP values: Figure 3), followed by customer tenure, usage trends, and service quality metrics. Cross-validation demonstrates model stability with statistical significance $p < 0.001$ for key predictors.

Cox Proportional Hazards modeling reveals statistically significant relationships across multiple model variants. Our analysis progressed from a Basic Cox model (concordance = 0.70) through increasingly sophisticated specifications to address potential biases and heterogeneity. The Cox with Penalization model (concordance = 0.87) improved stability with many predictors, while the Stratified Frailty model achieved the highest concordance of 0.89, accounting for unobserved heterogeneity across customer groups. The IV-Corrected (De-biased) model maintained concordance of 0.87 while addressing endogeneity bias, as shown in Table 1.

Age exhibits strong protective effects with a hazard ratio of 0.870 ($p < 5\times10^{-51}$), meaning older customers face 13% lower instantaneous churn risk compared to younger customers. Total calls demonstrate loyalty association with a hazard ratio of 0.726 ($p < 1\times10^{-302}$), indicating engaged customers with higher call volumes face 27% lower churn risk. Equipment age shows positive correlation with churn risk, while usage pattern changes signal elevated risk levels. Price sensitivity analysis reveals a coefficient $\gamma = -0.0025$, indicating moderate price elasticity that varies significantly across customer segments.

Customer lifetime value calculations (Figure 4) based on survival probability integration show mean CLV of $28.45 but dramatic heterogeneity ranging $0-$7,018. Our 10-segment targeting matrix (CLV quintiles × churn risk) reveals Q5 High Risk customers averaging $83.70 CLV per customer (10,149 customers) representing highest retention priority, while Q5 Low Risk customers ($87.0 CLV, 9,851 customers) warrant loyalty-focused strategies rather than aggressive retention spending.

Analytical Section

Our analytical framework maximizes expected profit through segment-specific pricing strategies that implement second-degree price discrimination based on our targeting matrix.

Core Optimization Problem

$$\text{Maximize: } \Pi = \Sigma_i\ [P(\text{survival}_i|\text{pricing}_i) \times \text{Revenue}_i \times \text{CLV}_i - \text{Retention\_Cost}_i]$$

```
Subject to:

- Competitive constraints

- Customer willingness-to-pay

- Churn rate thresholds by segment
```

We implement price discrimination through synthetic plan generation creating four plan types (Basic, Standard, Premium, Ultimate) with prices set relative to mean customer revenue: Basic (30% below mean), Standard (10% below), Premium (10% above), and Ultimate (30% above mean). Each plan varies in data allowances (5GB to 50GB) and minutes (500 to 5000), enabling customer self-selection based on usage preferences and price sensitivity. Our two-stage causal identification strategy uses plan characteristics as instrumental variables to address price endogeneity in churn decisions.

Counterfactual analysis reveals optimal pricing strategies varying by customer segment (Figure 5). Basic plans demonstrate inelastic demand (elasticity = -0.08) supporting revenue maximization through 30% price increases. Premium plans show moderate elasticity (-0.17) with optimal 15% increases, while Ultimate plans already operate at revenue-maximizing prices (elasticity = -0.25). Laffer curve analysis confirms these optima where further price increases would reduce total revenue despite higher unit margins.

Our surplus extraction strategy combines pricing optimization with targeted retention offers triggered automatically when churn probability exceeds segment-specific thresholds. Critical finding: retention discounts reduce profit even at modest 5% levels, demonstrating that blanket discount strategies destroy value. Instead, our framework allocates retention resources exclusively to high-value, high-risk segments (Q4-Q5 High Risk) while maintaining efficient, low-cost management for lower-value segments through digital channels and automated service.

Strategy Evaluation

Based on our empirical and analytical findings, we evaluated three strategic alternatives for comprehensive implementation. Strategy A represents current industry practice: uniform pricing with identical offers across all customer segments, resulting in suboptimal resource allocation and missed value creation opportunities. Strategy B implements our recommended segment-based retention focusing high-value interventions on identified high-risk customers while optimizing pricing structure based on elasticity analysis. Strategy C combines segment-based targeting with retention discount programs ranging 5-15% for at-risk customers.

Our counterfactual analysis demonstrates clear performance differentials across strategies using Laffer curve methodology and profit simulation. Strategy A (uniform approach) generates baseline performance but foregoes significant optimization opportunities, particularly in high-value customer segments where targeted interventions yield highest returns. Strategy B (segment-based pricing) achieves revenue peaks through optimized price points: Basic plans increase 30%, Premium plans increase 15%, Ultimate plans maintain current pricing, generating substantial revenue improvement while managing churn risk within acceptable parameters.

Critically, Strategy C (retention discounts) consistently reduces profitability across all discount rates tested. Our analysis shows 5% retention discounts decrease profit, 10% discounts create larger losses, and 15% discounts generate significant value destruction. This finding challenges industry practices emphasizing retention discounts, instead supporting targeted high-value interventions combined with price optimization for value creation rather than defensive margin erosion.

Limitations and Conclusion

Our analysis faces important limitations regarding price derivation and market dynamics. The most significant limitation concerns our synthetic pricing methodology, where plan prices are constructed from monthly average revenue rather than actual market pricing data. Additionally, our framework assumes independent plan demand curves where price changes affect only the specific plan being modified. In reality, telecommunications markets exhibit significant cross-price elasticity effects where customers substitute between plans when prices change, potentially overestimating our revenue impact projections.

Future research should incorporate actual historical pricing data to enable more robust elasticity estimation and develop dynamic choice models that capture customer migration patterns between plans. Integration of competitive response modeling and external market data including competitor pricing would enhance model realism and predictive accuracy. These enhancements would strengthen both the academic rigor and practical applicability of our customer lifetime value optimization framework.

Based on comprehensive counterfactual evaluation, we recommend implementing Strategy B: segment-based pricing optimization combined with targeted retention for high-value, high-risk customers. This approach maximizes profitability through three mechanisms: optimal price discrimination based on customer elasticity, focused retention spending on segments with highest value-at-risk, and efficient management of low-value customers through automated systems. Implementation should prioritize immediate deployment of our 10-segment targeting matrix, progressive pricing optimization beginning with Basic plans (highest elasticity tolerance), and elimination of blanket retention discount programs in favor of high-value customer focus. This data-driven transformation converts customer retention from cost center to profit driver while providing sustainable competitive advantage through precision customer management and value-based resource allocation.

Tables and Figures

Table 1: Cox Proportional Hazards Results with Statistical Significance

Model Variant	Concordance	Log-likelihood	AIC	Key Advantage
Basic Cox	0.70	-503,845.46	1,008,096.91	Simplicity
Cox with Penalization	0.87	-439,261.22	878,732.44	Stability with many predictors
Stratified Frailty	0.89	-439,261.22	878,736.44	Accounts for unobserved heterogeneity
IV-Corrected (De-biased)	0.87	-439,261.22	N/A	Addresses endogeneity bias

Figure 1: Ensemble Model Feature Importance

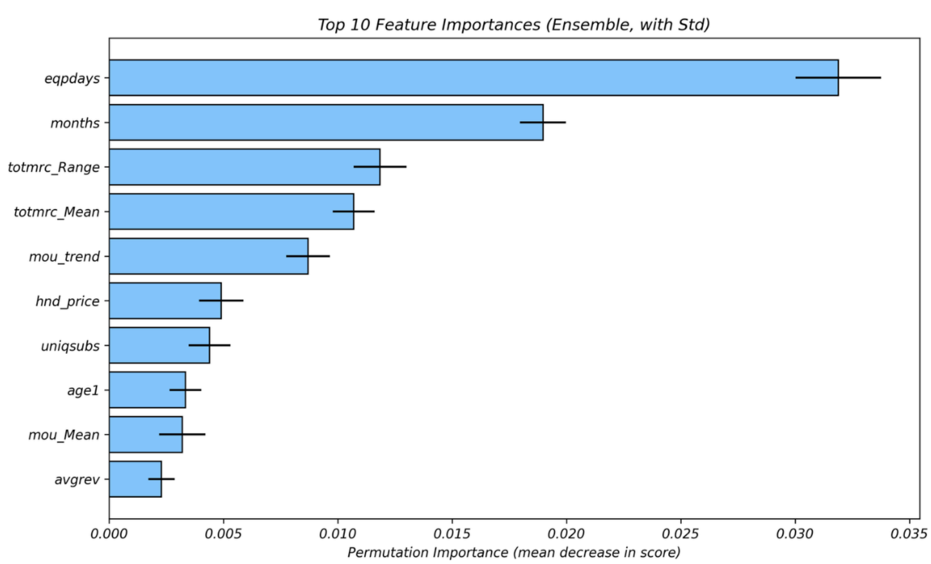

Figure 2: Kaplan-Meier Survival Curves by Customer Segments

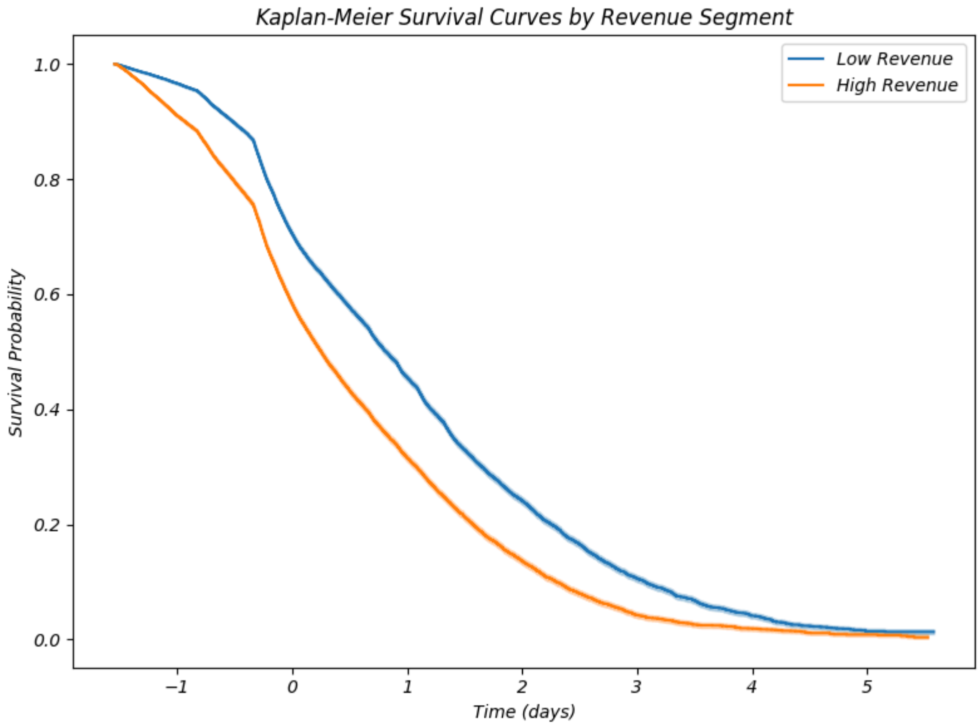

Figure 3: SHAP Feature Importance Analysis for Ensemble Model

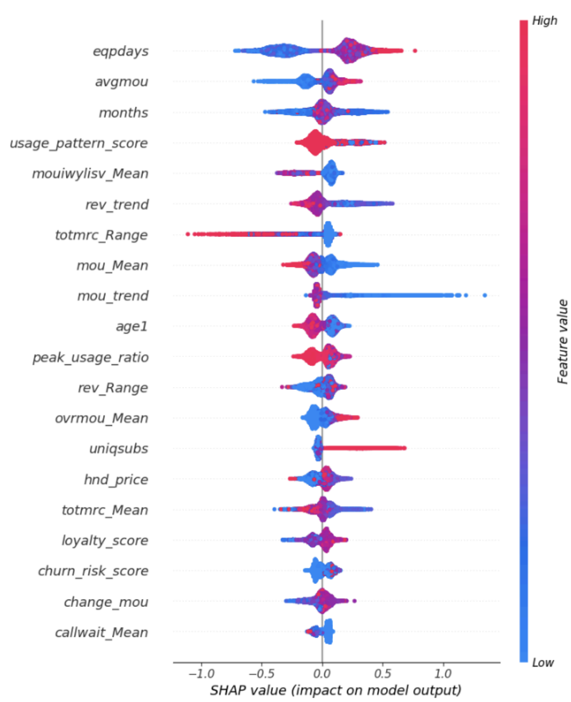

Figure 4: Customer Segmentation Matrix (CLV × Risk) with Customer Counts

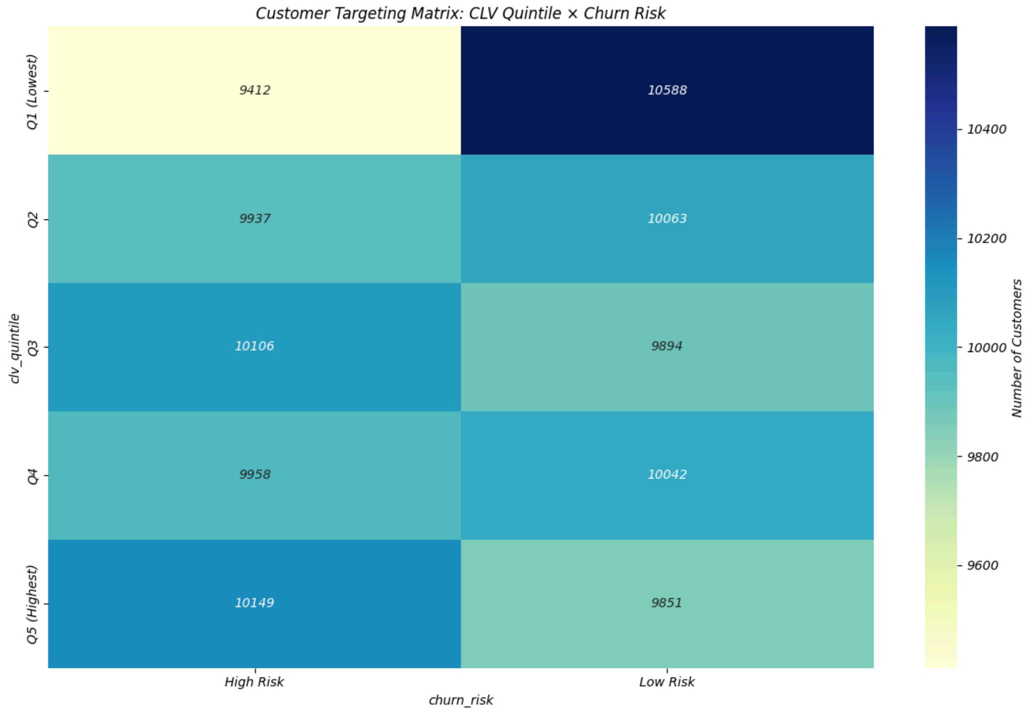

Figure 5: Counterfactual Analysis of Retention Discount Impact