A Novel Approach to Predicting Customer Lifetime Value in B2B SaaS Companies

Stephan Curiskis, Xiaojing Dong, Fan Jiang, Mark Scarr

(*Received 6 August 2024*)

Abstract

This report presents a flexible machine learning framework to predict customer lifetime value (CLV) in the B2B SaaS context. The framework addresses challenges related to customer relationships, heterogeneous populations, multiple product offerings, and constrained temporal data. It includes a hierarchical ensembled CLV model that integrates a variety of supervised learning techniques. The proposed method shows significant improvement over conventional forecasting methods.

1 Introduction

Understanding Customer Lifetime Value (CLV) is critical for B2B SaaS companies due to longer sales cycles and higher acquisition costs. This report explores a novel approach to CLV prediction, framing it as a lump sum prediction problem and utilizing a hierarchical ensembled model to address data constraints and customer heterogeneity.

2 Methodology

2.1 Problem Framing

The CLV estimation is treated as a lump sum prediction across multiple products. This allows the use of diverse supervised learning techniques, enhancing flexibility and feature richness. The hierarchical approach is particularly suitable for constrained temporal data, with a customer segment model ensembling strategy introduced for hyperparameter tuning.

2.2 Hierarchical T-Period Model

The hierarchical model involves two stages:

1. Training a T' period model using n periods of historical data.
2. Mapping T' period predictions to T period predictions using a second model that relies on slowly changing features such as firmographics.

Figure 1 illustrates this hierarchical approach.

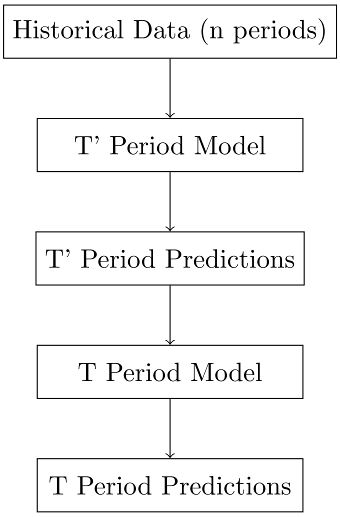

Fig. 1. Hierarchical T-Period Model for CLV Prediction

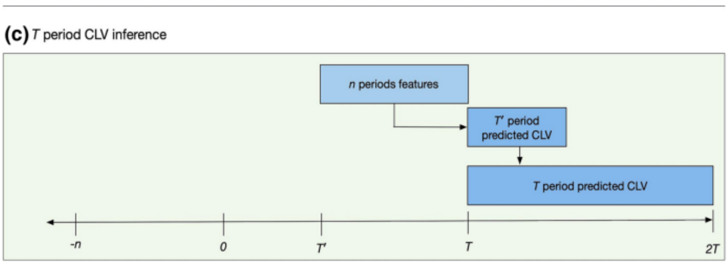

Fig. 2. Scatter points and regression line

3 Ensembled Customer Segment Model

Given the wide variation in CLV drivers, an ensembled approach is adopted. The data is segmented based on key features identified through error diagnostics, and different types of prediction models are applied to these segments.

4 Empirical Data and Results

The model was implemented on data from a B2B SaaS provider, with the T' period set to 2 years and the T period set to 5 years. The features included revenue trends, product usage data, and firmographic details.

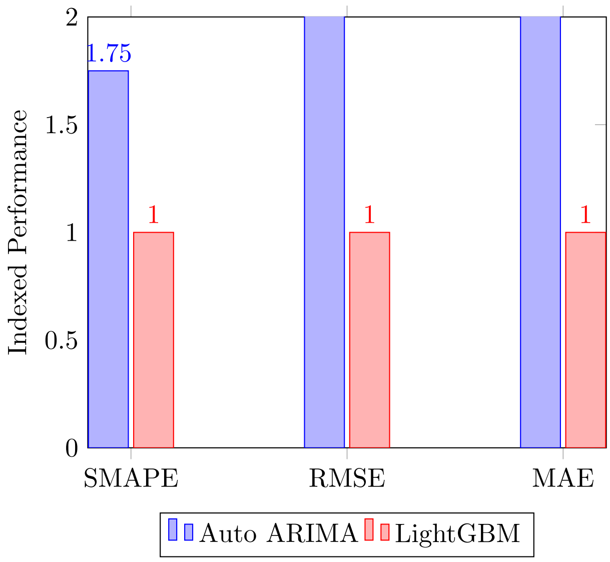

Fig. 3. Performance Comparison of Different CLV Prediction Models

5 Business Applications

5.1 Projected Value

The projected value is derived from the acquisition CLV, accounting for initial product revenue and future expansions. This metric supports budget planning and marketing ROI optimization.

$$\text{Projected Value} = \text{signups} \times \text{purchase rate} \times \text{acquisition CLV} \tag{1}$$

5.2 Return on Investment (ROI) Optimization

The ROI optimization framework uses CLV to guide marketing spend decisions, enhancing budget efficiency and targeting strategies.

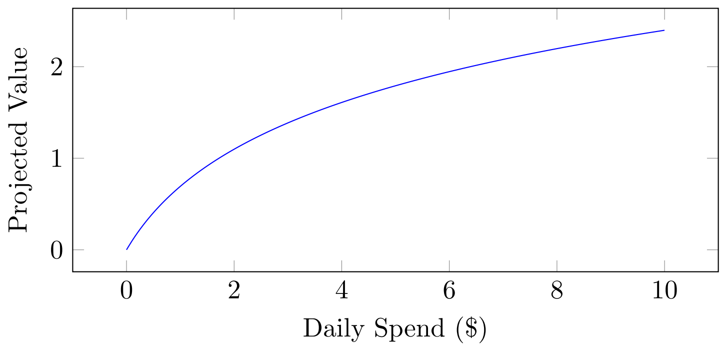

Fig. 4. ROI Optimization Framework

6 Conclusion

The proposed hierarchical ensembled CLV model addresses key challenges in B2B SaaS settings, offering significant improvements in prediction accuracy. This framework is generalizable to other contexts with similar challenges and can drive critical business insights for marketing, customer retention, and resource allocation.