International Journal of Engineering Applied Sciences and Technology, 2024
Vol. 9, Issue 03, ISSN No. 2455-2143, Pages 160-164
Published Online July 2024 in IJEAST (http://www.ijeast.com)

CUSTOMER SEGMENTATION AND LIFETIME VALUE PREDICTION

Sugat Bhagat, Kishor Deshmukh, Kajal Bonde, Rameshwari Fegade, Kshitija Dale
Computer engineering
Vishwakarma institute of technology
Pune,India

Abstract — **Customer segmentation and lifetime value estimation are two areas of development today. Today, the business world works differently, so machine learning uses different algorithms and different models to increase business efficiency. The value of life measured by its results is the bond between the customer and the company, as well as the value of the customer for the company. The reason for choosing this project is to examine the predictive behavior of future personalization models. From the company's perspective, this model can be very effective in terms of customer relations. Both segmentation and LVP are good for strategy development in business. As the world continues to evolve with major changes in technology, it is only natural to do such a project. LVP has a lot of predictive power when evaluating the customer buying process and choosing the best services to offer to different customers.**

Current methods for estimating CLV involve developing a model that uses all customer-based inputs. This not only loses the granularity of the data, but also makes the marketing plan less effective from the customer's perspective. This article focuses on the advantages of intelligently combining small models by focusing on separate models for different customers. This allows companies to effectively use customer data to make informed decisions. Using the data as an example, the proposed multi-project study where the learning experience is shared with relevant groups increases the CLV estimation accuracy compared to using a single model size with some minor modifications. Moreover, the same method reduces the standard deviation of the error when compared with a large sample. Most importantly, when the groups have different data to train, multiple learning models will outperform single-shot models, which is normally considered a more difficult task. These results show that multi-task learning can be more effective than existing business models and can be a better alternative to existing models.

Keywords—**clusters, K-means, lifetime value, RFM, segmentation**

I. INTRODUCTION

Market segmentation is the process of identifying key groups with specific characteristics and behaviors in the overall market, allowing products or services to be managed to meet their needs. Currently, the RFM model proposed by Hughes (1994) is one of the most widely used methods in the business world to segment and analyze the value of customers. This approach is based on the nature, frequency and results of the measurement, which are considered as three important factors to reveal customer behavior and influence their future purchases. Using the RFM model, marketing managers can target valuable customers and then develop marketing strategies for them according to their value. The purpose of this article is to propose a Lifetime Value Estimation (LVP) model and customer segmentation, considering the data used to create the model and the algorithm used. The flow, which starts with a description of some tasks and some ideas to arrive at the customer lifetime value estimate and their future reasons. Our model contains input data from online stores where many customers have different identities, different products purchased from different sources, and different prices.

In today's highly competitive environment, business people are under pressure to be more responsible in their business activities. Recent research on customer lifetime value (CLV) provides an important framework that clarifies marketing strategy for financial measurement. The framework evaluates how changes in customer behavior affect future customer outcomes or their benefits to the company. Although there are many recommendation methods, few consider customer lifetime value (CLV) and influence product recommendations. Businesses are increasingly aware of the importance of customer lifetime value. RFM (Recency, Frequency, and Monetary) method is generally used to measure CLV (Miglautsch, 2000; Kahan, 1998). In a competitive environment, determining the CLV or loyalty ranking of customers is important to help people make clear business decisions. Furthermore, the effect of CLV on referrals should be investigated to develop better marketing strategies. Recently, Liu and Shih (2004) proposed a weighted RFM-based method (WRFM-based method) that integrates data mining to recommend products according to customer needs. The WRFM-based approach uses organizational policy mining

to determine policies recommended by consumer groups based on weighted RFM values. Their experiments show that the WRFM-based method outperforms the KNN-based CF method and can determine good rules to provide useful customer recommendations that are good or trustworthy in life.
However, it is difficult to create policies that agree with less loyal customers. Similar to the WRFM-based method, the preference-based CF method can also be used by the mining authority to extract recommended rules from user groups based on their shopping preferences. The experiments of this research show that the preference-based CF method can provide some recommendations that the WRFM-based method cannot, thus improving the recommended message for less loyal customers. Therefore, this research leads to mixed results.'

II. OVERVIEW

Many studies have been done and published in this field. This study proposes to develop a rough system to check whether it will be useful for business management. This section only explores the problems or theoretical part of the operational and management model and also some ideas for the customer value group.
1.Once the number of clusters was identified, a k-means clustering algorithm, which is a non-hierarchical method, was used. These analyses at the end provided further illustrations of using cluster method for market segmentation for forecasting.
2. Random Forest model framework, seasonality and decay models. The lessons learned in building a customer lifetime value monitoring system for a large global ecommerce company.
3. Many CRM researches pertain to develop a comprehensive model of customer profitability since the question 'Who are profitable customers?' is a starting point of CRM. Many models have been researched to calculate LTV of a customer. Most of them focused on the future cash flow derived from the past profit contribution.
4. Interpurchase Time Model, Purchase-Amount Model, Customer-Defection Model, NBD–Pareto model.
5. Self-organizing maps (SOM) of neural network used for clustering and data visualization, SOM represent a good way of analyzing quantitatively the company's customer database no matter how big it is.
6. WEKA, J48 decision tree, RF (random forest), multilayer perceptron (MLP), and sequential minimal optimization (SMO) classifiers. For predicting customer defection, each classifier has its best criteria. The J48 Decision Tree and SVM models were excellent due to compatibility with the data sets.
7. SNA based algorithm, k means clustering and segmentation using groups.
8. Synthetic Minority Oversampling with Nominal and Numerical Attributes for Lifetime Value. In non-contractual freemium settings, users can freely choose the number and size of their purchases of premium upgrades, examples are freemium games and dating.
9. FCM clustering Customer segmentation, Customer relations, Customer lifetime value, Fuzzy AHP, Fuzzy c-means, Fuzzy logics used.

III. PROPOSE SYSTEM

K-means algorithm
Clustering is the process of grouping physical objects or data that are similar and different from objects in other groups into the same group. K-means is one of the well-known grouping algorithms, originally called the Forgy's method, and has been widely used in many fields such as data mining, data analysis, etc. Therefore, this work introduces a K-means algorithm to generate clusters based on behavior. The K-means algorithm used for classification is based on the mean of the objects in the cluster. MacQueen suggests using the term K-means to describe an algorithm that assigns each object to the cluster with the nearest center point (which we usually call the mean). Based on the above points, the calculation process of K-means is as follows:

Step 1: Partition the items into K initial clusters. Firstly, partition the items (m objects) into K initial clusters.
Step 2: Continue viewing the list. Assign an object with the closest centroid to the cluster (the distance is calculated using the Euclidean distance of the observation sample or without using the sample) and recalculate the centroid of the new products or missing products of the resulting cluster.
Step 3: Repeat Step 2 until no more reassigning. Rather than starting with a partition of all items into K preliminary groups in Step 1, we could specify K initial centroids (seed points) and then proceed to Step 2. The final assignment of items to clusters will be, to some extent, dependent upon the initial partition or the initial selection of seed points. Experience suggests that most major changes in assignment occur with the first reallocation step.

RFM model
The RFM analytic model is proposed by Hughes (1994), and it is a model that differentiates important customers from large data by three variables (attributes), i.e., interval of customer consumption, frequency and money amount. The detail definitions of RFM model are described as follows:

(1). Recency of the last purchase (R).
R represents recency, which refers to the interval between the time that the latest consuming behavior happens and present. The shorter the interval is, the bigger R is.

(2) Frequency of the purchases (F).
F represents frequency, which refers to the number of transactions in a particular period, for example, two times of

one year, two times of one quarter or two times of one month. The many the frequency is, the bigger F is.
(3) Monetary value of the purchases (M).
M represents monetary, which refers to consumption money amount in a particular period. The much the monetary is, the bigger M is.

The Pareto/NBD and BG/NBD models

They predict future actions based on each customer's past purchasing behavior. Each customer needs 3 historical measurements (T), which is the time elapsed since the customer joined the company, the "frequency" x, which is the number of times the customer has made this transaction k times subsequently, and the "recency" tx, which is the time between the entry date and the last purchase date. The Pareto/NBD model is based on six assumptions:

1) Customers go through two stages in their "lifetime" with a specific firm: they are "alive" for some period of time, and then become permanently inactive.
2) While alive, the number of transactions made by a customer follows a Poisson process with transaction rate, an Improved BG/NBD Approach for Modeling Purchasing Behavior.
3) Heterogeneity in transaction rates across customers follows a gamma distribution.
4) A customer's unobserved "lifetime" of length τ is exponentially distributed with dropout rate μ.
5) Heterogeneity in dropout rates across customers follows a gamma distribution.
6) The transaction rate λ and the dropout rate μ vary independently across customers.

Assumptions 2 and 3 give the Pareto distribution, and assumptions 4 and 5 give the NBD model. Although Pareto/NBD is a well-known model, it is difficult to apply due to the difficulties in calculating the measurement parameters. To overcome this problem, an alternative model to the Pareto/NBD model is called BG/NBD. The second model has the same assumptions as the Pareto/NBD model when modeling variables. On the other hand, the BG/NBD model does not maintain assumptions 4 and 5 because it is assumed that customers follow the Beta geometric distribution of the customer life cycle and people do not use immediately after purchase. Different from Pareto. In the NBD model, it is assumed that customers can interact at any time. Since the two models yield similar results in many purchasing environments, BG/NBD can be considered an interesting alternative to Pareto/NBD in many applications.

Gamma-Gamma model

Our model of spend per transaction is based on the following three general assumptions:
• The monetary value of a customer's given transaction varies randomly around their average transaction value.
• Average transaction values vary across customers but do not vary over time for any given individual.
• The distribution of average transaction values across customers is independent of the transaction process.

IV. METHODOLOGY

The dataset is taken from a UCI Machine Learning Repository I.e. Online Retail Dataset. This Dataset include fields such as Invoice No, Stock Code, Description, Quantity, Invoice Date, Unit Price, Customer ID and Country. To understand methodology, it is important to be clear about the definitions of models used in this study.

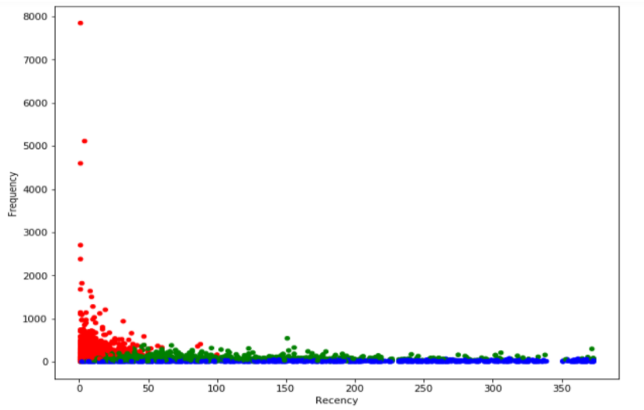
LTV = expected number of transactions * revenue per transaction * margin

Expected Number of Transactions were calculated by BG/NBD Model and revenue per transaction by Gamma-Gamma model.
Math notation for BG/NBD Model:
X = x, t_x, T, where x is the number of transactions at some period of time (0, T], and t_x (<=T) is the time of the last purchase.
Based only on these features, the model predicts future purchasing patterns of customers:
P(X(t) = x) => probability of observing x transactions in the period t in the future.
E(Y(t) | X = x, t_x, T) => expected number of transactions in the period for a customer with observed behavior.

Math notation for Gamma-Gamma model :
The customer has x transactions with z1, z2, … values,
m_x = Zi/x is observed mean transaction value.
E(M) is unobserved mean transaction value,
E(M | mx, x) - expected monetary value of a customer giving his purchasing behavior.

V. RESULT

After implementation first we get recency, frequency and monetary values of each customer i.e. we got RFM Score.
With the help of k means we got clusters points.



From BG/NBD model we get Expected Number of Transactions (pred_num_txn)

	CustomerID	frequency	recency	T	monetary_value	pred_num_txn
0	12748.0	113.0	373.0	373.0	298.360885	2.57
1	17841.0	111.0	372.0	373.0	364.452162	2.53
2	15311.0	89.0	373.0	373.0	677.729438	2.03
3	14606.0	88.0	372.0	373.0	135.890114	2.01
4	12971.0	70.0	369.0	372.0	159.211286	1.60
5	13089.0	65.0	367.0	369.0	893.714308	1.50
6	14527.0	53.0	367.0	369.0	155.016415	1.23
7	13798.0	52.0	371.0	372.0	706.650962	1.20
8	16422.0	47.0	352.0	369.0	702.472340	1.09
9	14096.0	16.0	97.0	101.0	4071.434375	0.99

From Gamma-Gamma model we calculated revenue per transaction. With the help of these we calculated customers Lifetime value (CLV)

	CustomerID	CLV
0	18102.0	177369.90
1	16446.0	162411.47
2	17450.0	146253.47
3	14096.0	125732.90
4	17511.0	67134.25
5	16029.0	58295.19
6	16684.0	50744.58
7	13694.0	48580.94
8	15311.0	46074.11
9	13089.0	44700.00

VI. CONCLUSION

This study presents a method that combines the features of RFM and K-means methods to improve the classification accuracy while eliminating outliers to achieve good results. It can also improve some shortcomings of the data mining tool. In order to demonstrate the methodology, online sales data (including 401 events) collected by the company is used as test data in this study. As can be seen the proposed system has met the specified criteria in terms of accuracy regardless of outputs the process is said to understand the decision. The result of the proposed process is an easy-to-understand decision-making process that makes it easier for companies to understand and know which customer is more important and which customers contribute the most to the company's revenue.

In addition, this method based on RFM behavior and K-means algorithm helps companies analyze their customers. Based on these positive test results, it is believed that this study will help companies focus on the customer's business, thus obtaining the maximum benefit and achieving a win-win between the company and the customer. Based on the results of this empirical study, we conclude that the proposed method is more effective than the specified method in distributing customer cost to RFM features, K-means algorithm and our model. For future research, other data types such as financial markets or healthcare markets or even service markets can be considered to evaluate this process, or other customer value measures (other than RFM model) can be used as features to distribute customers. In general, we hope that the proposed method is general for the entire dataset rather than exclusive.

VII. LIMITS AND FUTURE SCOPE

In this article, we will not discuss the factors that make customers loyal or the reasons why customers continue to buy a particular brand. Previous studies have examined various determinants of customer loyalty, such as customer satisfaction, trust, brand equity, social products, and purchase. We focus only on the relationship between two customer metrics: customer loyalty and estimated lifetime value. The researchers used aggregate data and multivariate models to examine the effects of service quality, customer retention, and customer satisfaction on equity. Unlike LVP, customer loyalty is a macro-level indicator that can be used directly to understand how a company responds to marketing campaigns. Our analytical framework can be used as a basis for calculating customer loyalty. To do this, researchers need product exposure data to predict future purchases. Extending our work to loyal customers is an opportunity for future research.

ACKNOWLEDGMENT

We would like to express my deep and sincere gratitude to our research supervisor Sangita Lade mam, to our college Vishwakarma Institute of Technology, Pune for having a course subject like Engineering Design and Innovation. During the project we came across different challenges which even led us to some unique concepts during the whole process, found some solution too. We are grateful for giving us the opportunity to do research and providing invaluable guidance throughout this research. Her dynamism, vision, sincerity and motivation have deeply inspired us all. She has taught us the methodology to carry out the research and to present the research works as clearly as possible. It was a great privilege and honor to work and study under her guidance. We are glad to ourselves for the keen interest shown to complete this thesis successfully.

VIII. REFERENCES

[1]. Velu C. M, 2013, Customer Segmentation Using Clustering and Data Mining Techniques, DOI:10.7763/IJCTE. 2013.V5.811

[2]. Pandey Addhyan, Han Angela and Parekh Rajesh, 2016, An Engagement-Based Customer Lifetime Value System for E-commerce, DOI: http://dx.doi.org/10.1145/2939672.2939693

[3]. Hwang Hyunseok, Jung Taesoo and Suh Euiho, 2016, An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry, https://doi.org/10.1016/S0957-4174(03)00133-7

[4]. Sifa Rafet, Runge Julian, Bauckhage Christian and Klapper Daniel, 2018, Customer Lifetime Value Prediction in Non-Contractual Freemium Settings: Chasing High-Value Users Using Deep Neural Networks and SMOTE, DOI:10.24251/HICSS.2018.115

[5]. AMPERITY INC, 2019, Predicting Customer Lifetime Value with Unified Customer Data.

[6]. Paramita Pradnya, Surjandri Isti, 2019 Estimating Customer Segmentation based on CustomerLifetime Value Using Two-Stage Clustering Method, DOI:10.1109/ICSSSM.2019.8887704

[7]. Daoud Rachid, 2015, Customer Segmentation Model in E-commerce Using Clustering Techniques and LRFM Model,

[8]. Prasasti Niken and Ohwada Hayato, 2014, Applicability of Machine-Learning Techniques in Predicting Customer Defection, DOI: 10.1109/ISTMET.2014.6936498

[9]. Rehman Ahsan, Raza Abbas, 2014, Customer Churn Prediction, Segmentation and Fraud Detection in Telecommunication Industry, Published by Semantic scholar, Corpus ID: 167500137

[10]. Singh Siddharth, Borle Sharad, 2007, A generalized framework for estimating customer lifetime value when customer lifetimes are not observed, DOI:10.1007/s11129-009-9065-0

[11]. Burelli Paolo, 2022, Predicting Customer Lifetime Value in Free-to Play Games, DOI:10.48550/arXiv.2209.12619

[12]. Shih Ya-Yueh and Liu Chung-Yuan, 2003, A method for customer lifetime value ranking - Combining the analytic hierarchy process and clustering analysis, DOI:10.1057/palgrave.dbm.3240216