Modelling Customers Lifetime Value For Non-Contractual Business

Riyanto [1,*], Abdul Azis [2]

[1,2] Universitas Amikom Purwokerto, Purwokerto
[1,*] riyanto@amikompurwokerto.ac.id; [2] abdazis9@amikompurwokerto.ac.id
* corresponding author

Abstract

Because customer equity is becoming increasingly important in today's business environment, many companies are focusing on customer loyalty and profitability as a way to grow market share. A company's successful Customer Relationship Management (CRM) begins with identifying true value and customer loyalty, as customer value can provide basic information that can be used to spread more targeted and personalized marketing. Customer lifetime value (CLV) is used in this document to segment consumers in non-contracted companies. The findings of this research are very promising. CLV has successfully analyzed and produced a fairly strong assumption about the value possessed by each customer whether they will make a return transaction or not.

Keywords: CLV, CRM, Business, E-Commerce, Customer Transaction Prediction

1. Introduction

Companies must build innovation activities to capture consumer desires and improve customer loyalty and retention in today's market, which is increasingly more dynamic and competitive [1]. In this way, customer relationship management is a well-known method for attracting and retaining clients. CRM's key aim is to develop long-term and successful customer partnerships [2]. A broad database containing comprehensive data on demographic details and consumer purchases is accessible in this context. To evaluate consumer equity, different CRM tools may be used to analyze this data. Customer Lifetime Value (CLV) is a CRM term that reflects the present value of all potential income produced by a customer [3]. CLV estimation has many uses, and many writers have developed models for them, including performance assessment [4], consumer segmentation [5], marketing capital distribution [6,7], product offering [8], pricing [9], and product offering. [10 - 12]

The relationship between the company and the buyer in E-Commerce or retail is a non-contractual relationship. Customers quit in the non-contract universe, but they do it quietly; they didn't have to tell us they were going. Calculating the CLV gets even more complex as a result of this. We would consider the period of time after the customer's last transaction to determine if the customer is alive but inactive or "dead" ("on" means the customer interacted with us, "dead" means they have become inactive as a customer). We will model the consumer lifetime for non-contracted companies in this report.

2. The Proposed Method/Algorithm

2.1. Customer Lifetime Value

CLV changed its focus away from the Customer Relationship Management issue (CRM). CRM is a company-wide technique for better understanding and shaping customer behaviour through positive dialogue in order to improve customer acquisition, retention, satisfaction, and profitability [13]. CRM's aim is to create closer and stronger relationships with consumers in order to increase their lifetime value to a brand [14]. There have been multiple classifications for the CLV model in previous research. Gupta et al suggested one of these divisions. Six modeling methods are described by Gupta et al: Recency, Frequency, and Monetary (RFM) model; A probability model based

on the Pareto / NBD model and the Markov chain; Customer acquisition, customer retention, and customer margins and expansion; econometric models such as chance models based on the Pareto / NBD model; The Longevity Model is focused on forecasting the behavior of its constituents, which involve acquisition, retention, and cross-selling. Models in computer science are theory-based (e.g., utility theory) and quick to grasp [15].

WRFM - weighted RFM - instead of RFM was suggested by several scholars in a recent study. R, F, and M each have their own range of weights. Different weights must be applied to the RFM parameters based on industry characteristics. For eg, Wei [16] suggests putting the most weight on Frequency, then Recency, and finally the monetary measure [18], but Chuang and Shen (2008) suggest putting the most weight on monetary and the least weight on highest recency [1]. The relative value (weight) of the RFM variables is calculated using the AHP equation.

2.2. Data Mining

Data mining is the method of automatically locating valuable information in large data warehouses. Data mining methods are used to sift through vast datasets in search of new and valuable trends that would otherwise go unnoticed [19]. There are two types of data mining methods: analytical and predictive. Grouping is a descriptive process, whereas classification is a statistical method. The method of discovering a model (or function) that represents and differentiates data classes or principles with the intention of using the model to predict the type of objects whose class mark is unknown is known as classification [20]. Unlike classification and prediction, which look at data objects with class identifiers, grouping looks at data objects without them. The CLV for each section is determined using the k-means clustering approach in this article. K-means, originally known as the Forgy process [21], is a well-known clustering algorithm that has been commonly used in a number of fields, including data processing, computational data analysis, and other market applications.

3. Method

Before starting the modeling and prediction process, we will describe the data we will use for today's research. We use the Online Retail Dataset which is available and can be downloaded open source in the UCI Machine Learning Repository. Let's take a closer look at the features in the dataset first, Attached in figure 1 below is the structure of the dataset that we use along with some sample data in it, there are approximately 550,000 data in it, but not all of the data will be used for manufacturing (Training & Testing) model. Besides that, there are still data that need to be cleaned because they still have a NaN value which will have a quite fatal effect when used for training models.

InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
546658	21506	FANCY FONT BIRTHDAY CARD,	12	2011-03-15 14:21:00	0.42	15150.0	United Kingdom
541422	48111	DOORMAT 3 SMILEY CATS	1	2011-01-17 17:48:00	14.13	NaN	United Kingdom
576463	21025	SPACE FROG	1	2011-11-15 11:37:00	1.25	17974.0	United Kingdom
550341	22635	CHILDS BREAKFAST SET DOLLY GIRL	2	2011-04-17 15:53:00	9.95	13018.0	United Kingdom
536981	22158	3 HEARTS HANGING DECORATION RUSTIC	1	2010-12-03 14:26:00	2.95	14723.0	United Kingdom

Fig. 1. Dataset Structure & Sample

As we said before, we have some cleanup to do, then create a new data frame containing only CustomerID, InvoiceDate (timeless) and add a new column which is "sales":

```
        CustomerID  InvoiceDate   Sales
305226     15505.0   2011-08-18   99.60
250194     15549.0   2011-07-05   39.80
490449     16164.0   2011-11-22   17.34
407654     12370.0   2011-10-19   23.40
7517       13838.0   2010-12-05    9.95
4339
```

Fig. 2. New Dataframe

The following nomenclature is used for CLV models: Frequency refers to the amount of times a consumer has made the same transaction. That is, it is less for one of the overall sales price. T denotes the customer's age in the time unit chosen (daily, in our data set). This is the time interval between a customer's first transaction and the completion of the testing era. Recency suggests the customer's age at the time of their most recent order. This is the same as the time period between a customer's first and last transaction. (The recency is 0 if they just make one purchase.)

CustomerID	frequency	recency	T	monetary_value
12346.0	0.0	0.0	325.0	0.000000
12347.0	6.0	365.0	367.0	599.701667
12348.0	3.0	283.0	358.0	301.480000
12349.0	0.0	0.0	18.0	0.000000
12350.0	0.0	0.0	310.0	0.000000

Fig. 3. CLV Dataframe

Our database currently includes 4339 clients. CustomerID 12346, for example, made only one transaction (no repeat), so the frequency and recency are both zero, and the age is 325 days (for example, the duration between the first purchase and the end of the period in the analysis). More than 35% of all consumers in our database just made a single order (no repeat).
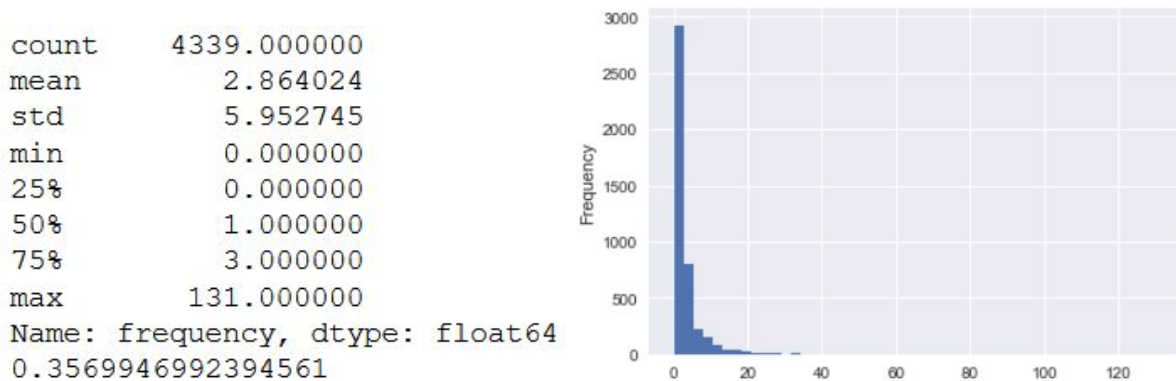
```
count      4339.000000
mean          2.864024
std           5.952745
min           0.000000
25%           0.000000
50%           1.000000
75%           3.000000
max         131.000000
Name: frequency, dtype: float64
0.3569946992394561
```

Fig. 4. Customer Purchase Frequency

```python
from lifetimes import BetaGeoFitter

bgf = BetaGeoFitter(penalizer_coef=0.0)
bgf.fit(data['frequency'], data['recency'], data['T'])
print(bgf)
```

```
<lifetimes.BetaGeoFitter: fitted with 4339 subjects, a: 0.00, alpha: 69.0
1, b: 1.18, r: 0.83>
```

Fig. 5. Frequency / Recency Analysis Using the BG / NBD Model

3.1. Visualizing our frequency/recency matrix

To begin, we must consider the fact that the customer has made a payment every day for the past four weeks, and yet we haven't heard from him in months. Is it true that he is still "alive"? Isn't it very small? Customers who have made sales once a year in the past and twice in the most recent quarter, on the other hand, are most definitely still alive. The frequency / recency matrix, which measures the estimated amount of purchases a consumer will make in the next

time frame depending on their recency (age at last purchase) and frequency, can be used to visualize this relationship (number of recurring transactions it has made).
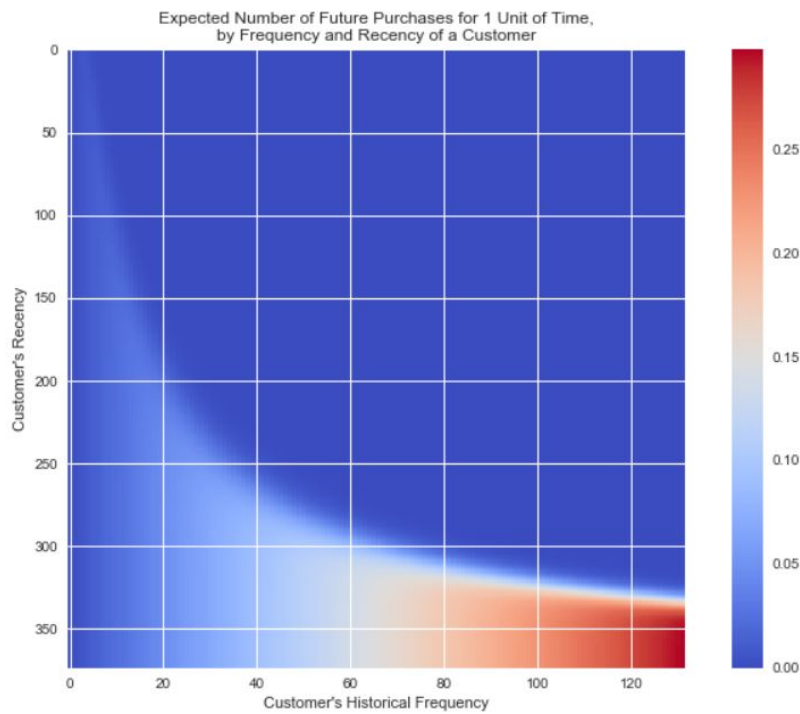

Fig. 6. Expected Number of Future Purchases for 1 Unit of time

When a client has made 120 transactions and his most recent purchase was 350 days ago (i.e., Recency: the time period between the first and last transaction was 350 days), he is our best customer (bottom right). A customer who has ordered regularly and lately is likely to be the best customer in the future. We'll never be able to get enough of them. Customers that have recently (top right corner) purchased a significant quantity of goods may have left. Other forms of customers (40, 300) reflect a client who seldom buys, but we haven't seen him in a while, so he'll most likely buy again. We're not sure whether he's gone or whether he just made a one-time buy. In the end we can predict which customer is definitely still alive:
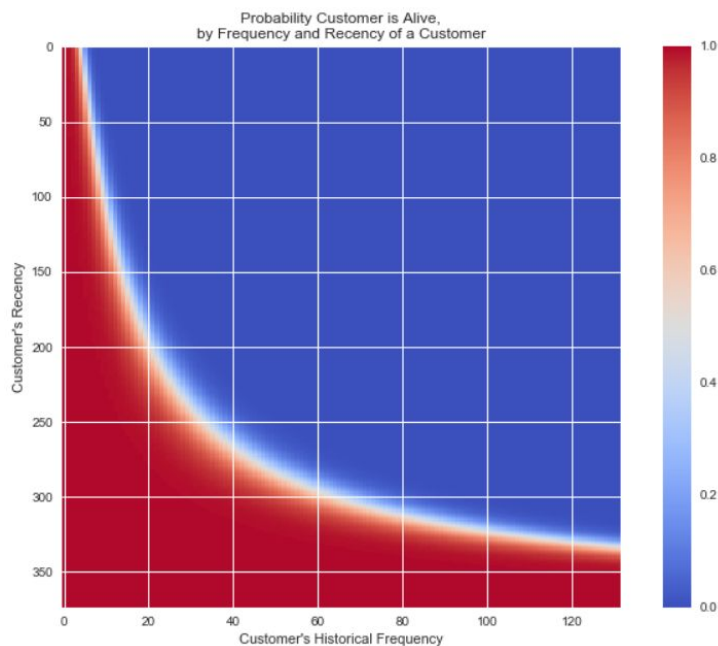

Fig. 7. Probability Customer is Alive

In figure 7 above, we can conclude that customers who have just bought must still be "alive". Customers who have bought a lot but not recently, have most likely left. And the more they bought in the past, the more likely they were to quit. They are represented at the top right. We rank customers from "highest estimated purchase of next period" to lowest. The model represents a method that will predict the purchases customers expect in the next period using their history.

CustomerID	frequency	recency	T	monetary_value	predicted_purchases
14606.0	88.0	372.0	373.0	135.890114	0.200966
15311.0	89.0	373.0	373.0	677.729438	0.203229
17841.0	111.0	372.0	373.0	364.452162	0.253003
12748.0	113.0	373.0	373.0	298.360885	0.257528
14911.0	131.0	372.0	373.0	1093.661679	0.298252

Fig. 8. Customers with the highest likelihood of returning (Based on the BG / NBD model)

Above (Figure 8) are our top 5 customers that the model is expecting them to buy the next day. The predict purchases column shows the number of purchases expected while the other three columns show their current RF metric. Model BG / NBD believes that these people will be making more purchases in the near future because they are our best customers at the moment.

3.2. Assessing Model Fit

The results we got were quite acceptable, the output model was also not too bad and usable. So, we can continue with our analysis. We now partition the dataset into a calibration period dataset and a split dataset. This is important because we want to test the performance of our model on unseen data (such as cross validation in machine learning practice).
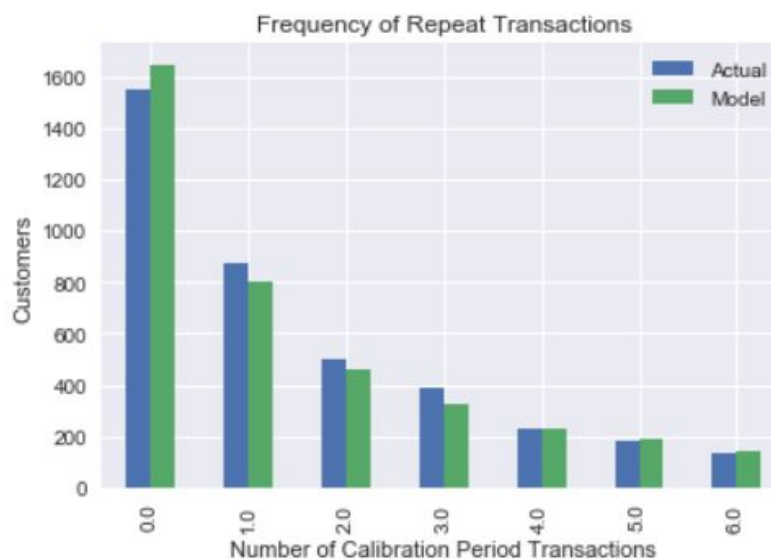

Fig. 9. Frequency of Repeat Transactions

```
              frequency_cal  recency_cal  T_cal  frequency_holdout  \
CustomerID
12346.0                0.0          0.0  141.0                0.0
12347.0                2.0        121.0  183.0                4.0
12348.0                2.0        110.0  174.0                1.0
12350.0                0.0          0.0  126.0                0.0
12352.0                3.0         34.0  112.0                3.0

              duration_holdout
CustomerID
12346.0                    184
12347.0                    184
12348.0                    184
12350.0                    184
12352.0                    184
```


Actual Purchases in Holdout Period vs Predicted Purchases
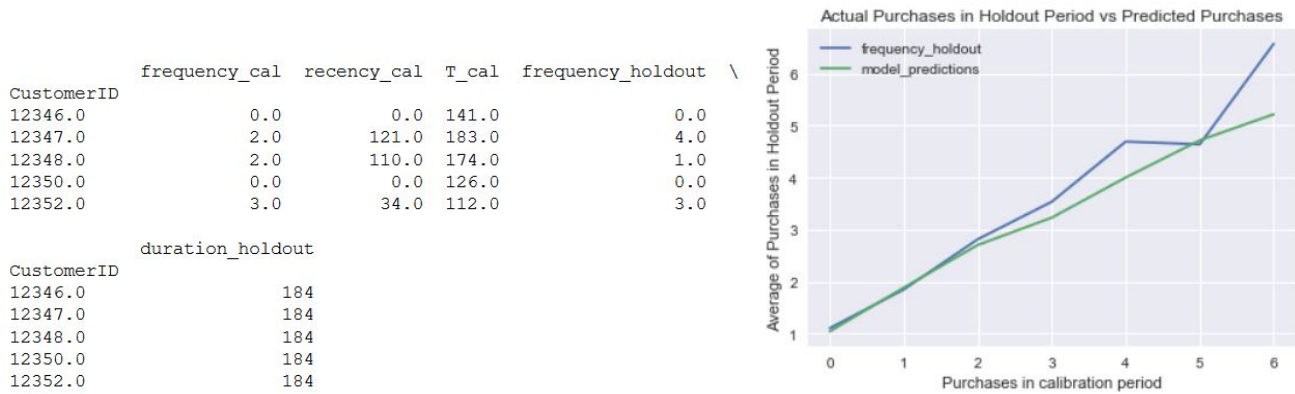
Fig. 10. Actual Purchases in Holdout Period vs Predicted Purchases

We divide the data in the plot above (Figure 10) into sample (calibration) and validation cycles (splits). The validity period runs from 2011–06–09 to 2011–12–09, while the survey period runs from early to 2011–06–08. The plot divides all consumers in the calibration cycle by the amount of repeated transactions they make (x-axis) and then averages those purchases over the split period (y-axis). The green and blue lines on the y-axis reflect the model prediction and actual outcome, respectively. As we can see, our model can predict the behavior of the out-of-sample, under-forecast consumer base at 4 and 5 transactions with considerable precision.

3.3. Customer Transaction Prediction & Probability Histories

```
t = 10
individual = data.loc[12347]
bgf.predict(t, individual['frequency'], individual['recency'], individual['T'])
```

0.1572775101289126

Fig. 11. Customer Transaction Prediction

Based on the customer's history, we can now predict what individual purchases will be in the future. Our model predicts that the future 12347 customer transactions will be 0.157 in 10 days. Based on a customer's transaction history, we can calculate their historical probability of staying alive, according to our trained model. For example, we want to see the transaction history of our best customers and see if they are still alive:

```
from lifetimes.plotting import plot_history_alive
import matplotlib.pyplot as plt

fig = plt.figure(figsize=(12,8))
id = 14606
days_since_birth = 365
sp_trans = df.loc[df['CustomerID'] == id]
plot_history_alive(bgf, days_since_birth, sp_trans, 'InvoiceDate')
```

Fig. 12. Model for Customer probability histories

Fig. 13. Customer probability histories model output

Our best customer is definitely still alive, however, he experienced several trading lags in the middle of 2011. On the other hand, this customer only bought once and has been down for a long time with a 100% chance of him dying.

4. Results and Conclusion

The current study focuses on customer segmentation as one of the CLV applications. As a case study, customer data from non-contractual businesses was examined. Using the Algorithm, we divide customers into segments based on RFM and Extended RFM parameters. Customer segmentation allows decision makers to more clearly identify market segments and develop more effective marketing and sales strategies for customer retention. The CLV method is used to determine the relative importance of the RFM variables based on the point of view of the expert in the sales department, because the RFM weights vary with industry characteristics. For each customer segment, the CLV value is calculated using the weighted RFM parameter. After that, each segment is given a CLV rating based on its CLV value. Potential value represents cross-selling opportunities, while present value provides a financial perspective. We can develop a refined marketing strategy for each segment by analyzing the CLV ratings of a segmented customer group. Our future work will be to implement this strategy in the company.

References

[1] C. C. Lin and D. H. Shih, "Data mining techniques to enhance customer lifetime value," *Adv. Mater. Res.*, vol. 225–226, pp. 3–7, 2011, doi: 10.4028/www.scientific.net/AMR.225-226.3.

[2] V. Ravi, *Advances in banking technology and management: Impacts of ICT and CRM*. 2007.

[3] T. L. Keiningham, L. Aksoy, L. Williams, and A. J. Buoye, "Customers as Assets," *Wallet Alloc. Rule*, no. October, pp. 89–108, 2015, doi: 10.1002/9781119209867.ch4.

[4] V. A. Z. Roland T. Rust, Katherine N. Lemon, "Return on Marketing:Using Customer Equity to FocusMarketing Strategy," *Journal of Marketing*, vol. 9992 LNAI. pp. 291–302, 2004, doi: 10.1007/978-3-319-50127-7_24.

[5] M. Haenlein, A. M. Kaplan, and D. Schoder, "Valuing the Real Option of," *J. Mark.*, vol. 70, no. July, pp. 5–20, 2006.

[6] W. Reinartz, J. S. Thomas, and V. Kumar, "Balancing acquisition and retention resources to maximize customer profitability," *J. Mark.*, vol. 69, no. 1, pp. 63–79, 2005, doi: 10.1509/jmkg.69.1.63.55511.

[7] M. Ma, Z. Li, and J. Chen, "Phase-type distribution of customer relationship with Markovian response and marketing expenditure decision on the customer lifetime value," *Eur. J. Oper. Res.*, vol. 187, no. 1, pp. 313–326, 2008, doi: 10.1016/j.ejor.2007.03.018.

[8] Y. Y. Shih and D. R. Liu, "Product recommendation approaches: Collaborative filtering via customer lifetime value and customer demands," *Expert Syst. Appl.*, vol. 35, no. 1–2, pp. 350–360, 2008, doi: 10.1016/j.eswa.2007.07.055.

[9] P. Hidalgo, E. Manzur, S. Olavarrieta, and P. Farías, "Customer retention and price matching: The AFPs case," *J. Bus. Res.*, vol. 61, no. 6, pp. 691–696, 2008, doi: 10.1016/j.jbusres.2007.06.046.

[10] R. Saharon, "Customer Lifetime Value Modeling and Its Use for Customer Retention Planning." Association for Computing Machinery, 2002.

[11] M. Haenlein, A. M. Kaplan, and A. J. Beeser, "A Model to Determine Customer Lifetime Value in a Retail Banking Context," *Eur. Manag. J.*, vol. 25, no. 3, pp. 221–234, 2007, doi: 10.1016/j.emj.2007.01.004.

[12] D. F. Benoit and D. Van den Poel, "Benefits of quantile regression for the analysis of customer lifetime value in a contractual setting: An application in financial services," *Expert Syst. Appl.*, vol. 36, no. 7, pp. 10475–10484, 2009, doi: 10.1016/j.eswa.2009.01.031.

[13] R. Swift, "How Leading Companies Achieve Profitability through Accelerating Customer Relationships," *Inf. Age Econ.*, no. September, pp. 35–42, 2001, doi: 10.1007/978-3-642-57547-1_5.

[14] D. Peppers, M. Rogers, and B. Dorf, "Is your company ready for one-to-one marketing?," *Harv. Bus. Rev.*, vol. 77, no. 1, pp. 151–160, 1999.

[15] P. Čermák, "Customer Profitability Analysis and Customer Life Time Value Models: Portfolio Analysis," *Procedia Econ. Financ.*, vol. 25, no. May, pp. 14–25, 2015, doi: 10.1016/s2212-5671(15)00708-x.

[16] J. Wei, S. Lin, and H. Wu, "A review of the application of RFM model," *African J. Bus. Manag.*, vol. 4, no. 19, pp. 4199–4206, 2010.

[17] C. H. Cheng and Y. S. Chen, "Classifying the segmentation of customer value via RFM model and RS theory," *Expert Syst. Appl.*, vol. 36, no. 3 PART 1, pp. 4176–4184, 2009, doi: 10.1016/j.eswa.2008.04.003.

[18] N. Piersma and J. J. Jonker, "Determining the optimal direct mailing frequency," *Eur. J. Oper. Res.*, vol. 158, no. 1, pp. 173–182, 2004, doi: 10.1016/S0377-2217(03)00349-7.

[19] M. Tarokh and M. EsmaeiliGookeh, "A New Model to Speculate CLV Based on Markov Chain Model," *J. Ind. Eng. Manag. Stud.*, vol. 4, no. 2, pp. 85–102, 2017, doi: 10.22116/jiems.2017.54609.

[20] P. S. Fader, B. G. S. Hardie, and K. A. L. O. K. Lee, "Rfm_Clv-Jmr-2005," vol. XLII, no. November, pp. 415–430, 2005.

[21] M. Tolon and D. Aslan, "Comparing Customer Segmentation With CLV Using Data Mining and Statistics: A Case Study," *J. Bus. Res. - Turk*, vol. 10, no. 4, pp. 887–900, 2018, doi: 10.20491/isarder.2018.554.