

An Empirical Assessment of Customer Lifetime Value Models within Data Mining

Abdulkadir HIZIROGLU¹, Merve SISCI², Halil Ibrahim CEBECI³,
Omer Faruk SEYMEN³

¹Bakircay University, The Campus, 35665, Izmir, Turkey,

²Dumlupinar University, The Central Campus, 43100, Kutahya, Turkey,

³Sakarya University, The Esentepe Campus, 54050, Sakarya, Turkey

kadir.hiziroglu@bakircay.edu.tr, sevkiyemerve.oge@dpu.edu.tr,
hcebeci@sakarya.edu.tr, ofseymen@sakarya.edu.tr

Abstract. Customer lifetime value has been of significant importance to marketing researchers and practitioners in specifying the importance level of each customer. By means of segmentation which could be carried out using value-based characteristics it is indeed possible to develop tailored strategies for customers. In fact, approaches like data mining can facilitate extraction of critical customer knowledge for enhanced decision making. Although the literature has several analytical lifetime value models, comparative assessment of the existing models especially within the context of data mining seems a missing component. The aim of this paper is to compare two different customer lifetime value models within data mining. The evaluation was carried out within the context of customer segmentation using a database of a company operating in retail sector. The results indicated that two models yield the same segmentation structure and no statistical differences detected on the select control variables. However, the remaining model produced rather different segmentation results than their peers and it was possible to identify the most lucrative model according to the statistical analyses that were carried out on the select control variables.

Keywords: customer lifetime value, customer segmentation, lifetime value modelling, data mining, customer analytics

1. Introduction

Customer lifetime value (CLV) modelling is an analytical component of customer relationship management and has been widely utilized by a variety of companies across different sectors including finance and insurance, retail and telecommunications in order to identify the differences between the customers. It is a measurement of a firm's net cash flows generated by its customers within specified lifetime duration (Gupta and Lehmann, 2003). Calculating lifetime value of customers precisely can help companies to position them and to differentiate the most appropriate services. There have been several lifetime value models in the related literature and these models can be classified

into two groups: past customer behavior models and future-past customer behavior models. There are mainly two differences between the past customer behaviour and the future-past customer behaviour models. The first difference is based on the assumption that whether the customers who are subject to assessments will be active or not in the future, while the second difference stems from the inclusion of costs of customers into the models. PCV Model (past customer value); RFM Model (recency, frequency, monetary); SOW Model (share of wallet) can be included in the first category which calculate the lifetime values by only using the past data of customers. As far as the second category of the models is concerned, although they all take the future behaviour of customers into consideration (Kumar, 2005), some analytical models (Berger and Nasr, 1998; Gelbrich and Wünschmann, 2007; Gupta and Lehmann, 2003; Ramakrishnan, 2006; Rust et al. 2004; Venkatesan and Kumar, 2004) include acquisition cost when calculating lifetime values while some others (Bauer et al., 2003; Bruhn, 2003; Collings and Baxter, 2005) do not so. The vast majority of the literature focuses on the latter category of the models either in modelling or empirical form, however, the current literature lacks of comparative research on evaluating those CLV models, especially within the context of segmentation (Lemon and Mark, 2006).

The aim of this paper is to make a comparison between two customer lifetime value models from segmentation perspective within data mining. The rest of the paper is organized as the followings. The empirical studies of the related literature are provided in Section 2. Section 3 presents the research method followed. Empirical research results, including calculation of lifetime values for each model and the segmentation structures obtained by the comparative models, and their assessments were presented in section 4. In the last section of the article, conclusions and recommendations from both academic and practical points were provided.

2. Literature Review

When the current literature on customer lifetime value modelling is examined the models can simply be classified into two groups: the models that take into account past customer behavior and the models consider both past and future behaviors. Every past customer behavior group models have unique parameters which is directly related to model's characteristics. Among the models RFM is most widely used one and it has been utilized in marketing areas for almost decades (Gupta et al., 2006). The future-past customer behavior models share the same principle that for every customer how long it will be active is determined then net present values of these customers are calculated throughout the activation period. Based on this principle most of the models use common variable/constant parameters such as retention rate, marketing cost, cash flow ratio and reduction rate.

Most of the studies on future-past customer behaviour models use retention rate to determine the activation period (Blattberg and Deighton, 1996; Berger and Nasr, 1998; Pfeifer and Carraway, 2000; Bauer et al., 2003; Bruhn, 2003; Gupta and Lehmann, 2003; Bejou et al., 2006; Kumar et al., 2008; Wiesel et al, 2008; Drèze and Bonfrer, 2009; Kumar and Shah; 2009). However, some of the models use different set of criteria such as loyalty (McDonald, 1996; Kim and Cha, 2002), number of purchase period (Dwyer, 1997), length of service (Rosset et al, 2003; Hwang et al., 2004; Gelbrich and

Wünschman, 2007), recent transaction time / recency (Chang and Tsay, 2004; Fader et al., 2004), frequency of buying (Chang and Tsay, 2004; Fader et al., 2004; Rust et al., 2004; Ramakrishnan, 2006). Within the activation period determination of the monetary values of all customers is crucial. So, almost every future-past customer behaviour models include a monetary-oriented variable. The most common variables in these models are; marketing cost (Berger and Nasr, 1998; Pfeifer and Carraway, 2000; Bauer et al., 2003; Bruhn, 2003; Gupta and Lehmann, 2003; Venkatesan and Kumar, 2004; Bejou et al., 2006; Gelbrich and Wünschman, 2007; Kumar et al., 2008; Drèze and Bonfrer, 2009), cash flow ratio (Dwyer, 1997; Berger and Nasr, 1998; Hoekstra and Huizingh, 1999; Pfeifer and Carraway, 2000; Bauer et al., 2003; Bruhn, 2003; Gupta and Lehmann, 2003; Fader et al., 2004; Kumar et al., 2008; Wiesel et al., 2008; Kumar and Shah, 2009) and reduction rate (Berger and Nasr, 1998; Jain and Singh, 2002; Bauer et al., 2003; Bruhn, 2003; Venkatesan and Kumar, 2004; Gupta and Lehmann, 2003; Rust et al., 2004; Collings and Baxter, 2005; Gelbrich and Wünschman, 2007). Also, different parameters and variables complement these monetary values like acquisition rate and cost (Blattberg and Deighton, 1996; Gupta et al., 2004), discount rate (Blattberg and Deighton, 1996; Rosset et al., 2003; Malthouse and Blattberg, 2005), purchase intention (Kim and Cha, 2002), monetary value (Chang and Tsay, 2004), expected revenue (Malthouse and Blattberg, 2005), contributed value (Aeron et al., 2008).

It is possible to find empirical studies in the related literature that utilized one of the past customer behaviour models. Most of the empirical studies use RFM models or its extensions. These studies use different datasets from different sectors such as retail (Liu and Shih, 2005a; Chen et al., 2009; Albadvi and Shahbazi, 2010; Chang and Tsai, 2011; Lin and Shih, 2011; Hu et al., 2013), Banking (Khajvand and Tarokh, 2011), textile (Golmah and Mirhashemi, 2012), wholesale (Chuang and Shen, 2008), healthcare (Khajvand et al., 2011) and charity organizations (Jonker et al., 2004). Some authors use well-known RFM extension called LRFM (or RFML) which include one or more parameters related to relationship length (or period of activity) (Hosseini et al., 2010; Lin et al., 2011; Alvandi et al., 2012; Parvaneh et al., 2012; Wu et al., 2014). Considerable amount of studies use different methods including generalized regression, logistic regression, quantile regression, latent class regression, CART, Markov chain modelling, neural network to create past customer behaviour model (Haenlein et al., 2007; Benoit and Poel, 2009; Hosseini and Tarokh, 2011; Chen and Fan, 2013; Ekinici et al., 2014).

Aforementioned future-past customer behaviour models were used in different empirical studies in the related literature too. Reinartz and Kumar (2000) utilized Berger and Nasr (1998)'s model in retail sector. The same model was also used in petroleum (Gloy et al., 1997) and banking (Glady et al., 2009) sectors. Hwang et al., (2004) designed a conceptual model and used it in their empirical study in telecommunications sector. Chen et al. (2009) used the same model in retail industry. Gupta et al. (2004) also utilized their own conceptual model with internet company datasets. Additionally Kim et al. (2006), Cuadros and Dominguez (2012), Guo et al. (2013) and Glady et al. (2015) used Kim and Kim (1999)'s basic structural model as well as Fader et al. (2004)'s and Fader et al. (2005)'s models. Wu and Li (2011) performed a CLV calculation using the models of McDonald (1996) and Kim and Cha (2002). Kumar et al. (2008) adapted three different CLV models that belong to Reinartz and Kumar (2000), Rust et al. (2004) and

Venkatesan and Kumar (2004) to perform an empirical study in information technology sector.

In recent years, customer analytics has attracted a great deal of attention from both researchers and practitioners. It is an indisputable issue that customer relationship management is a broad topic with many layers, one of which is data mining, and that data mining is a method or tool that can aid companies in their quest to become more customer-oriented. Data mining process uses a variety of data analysis and modeling techniques to discover patterns and relationships in data that are used to understand what your customers want and predict what they will do. Data mining can help companies to select the right prospects on whom to focus, offer the right additional products to company's existing customers and identify good customers who may be about to leave. Data mining can predict the profitability of prospects as they become active customers, how long they will be active customers, and how likely they are to leave. In addition, data mining can be used over a period of time to predict changes in details. For example, a firm could use data mining to predict the behavior surrounding a particular lifecycle event (e.g., retirement) and find other people in similar life stages and determine which customers are following similar behavior patterns.

The significance usage of data mining techniques provides advantages in the areas of modeling CLV, including performing analysis based on CLV and evaluating the optimal method for identifying customer lifetime value in many industries such as retail, insurance, banking, telecommunication, financial services (Kim et al., 2006; Liu and Shih, 2005a; Cheng and Chen, 2009; Chen et al., 2009; Azadnia et al., 2012; Khajvand and Tarokh, 2011; Lin et al., 2011; Alvandi et al., 2012; Parvaneh et al., 2012; Golmah and Mirhashemi; 2012; Chen and Fan, 2013; Hu et al., 2013). These techniques include decision tree, clustering, logistic regression, artificial neural network, support vector machine, random forests, survival analysis, association rule apriori and self-organising maps. On one hand, while modeling techniques provide capability of CLV estimation, companies have competitive advantages in terms of making decisions due to the analysis activities based on CLV via data mining. On the other hand, comparative studies present effectiveness of models for different cases in different situations. The customer lifecycle provides a good framework for applying data mining to CRM. On the "input" side of data mining, the customer lifecycle tells what information is available. On the "output" side, the customer lifecycle tells what is likely to be interesting (Freeman, 1999). Briefly, data mining has become an indispensable tool for both obtaining CLV and studies on CLV.

When the existing empirical studies are reviewed, there are many different models which either use past or past-future information to calculate CLV values. However, it is difficult to find a comparative study with regards to the evaluation of different lifetime value models from practical benefits and academic point of view, especially within the scope of data mining and segmentation. Lemon and Mark (2006) also highlighted this specific issue as they made a recommendation on comparing current lifetime value models from the perspective of the ability to generate more efficient segmentation structures. This paper contributes to the current literature by providing the results of an empirical work conducted on two different representative models, which are RFM and Gelbrich and Wünschmann Model, using a database in a comparative framework based on data mining methodology with a special focus on segmentation.

3. Methodology

Some of the previous empirical lifetime value studies that used large-scale customer data demonstrate the broad usage of data mining methodology for the lifetime value modelling problem and the usefulness of such methodology. The aim of this study is to compare two different customer lifetime value models within the context of customer segmentation. Based on the classification provided in the previous section two representative models from the groups of models were compared and an assessment using some control variables were carried out within segmentation context. In order to accomplish that the variables in the acquired databases were operationalised based on some assumptions for each model and they were put them in place to perform the analyses and the comparison. RFM model and Gelbrich and Wüschmann Model were chosen for comparison as they all need the same set of variables.

The dataset was procured from a supermarket retail chain in the UK that includes four consecutive months of around 300,000 customers. A simple random sampling methodology was employed to extract the research sample. As far as the size of the sample is concerned, approximately 1% of the database was used as the study sample. A sample of 3,017 was obtained for conducting the analyses. In order to perform the analyses, the author used a tailored software program code in C++ language. Also, IBM SPSS Modeler and Microsoft Excel were used to obtain descriptive and test statistics.

The dataset includes fields such as customer number, store ID, cashier ID, date of transaction, time of transaction, transaction value, number of unique products bought, total number of products bought and tender type. However, the data fields necessary to conduct the analyses were obtained. The operationalization of these variables for each model was provided in Table 1 and Table 2.

To understand methodology of the proposed comparison, it is important to be clear about the definitions of two models used in this study. By contrast to the other two models, RFM model is based on the past customer purchase behavior and R, F, M notations indicate Recency, Frequency and Monetary values, respectively.

Table 1. Operationalization of the Variables for RFM Model

Variable	Explanation	Operationalization
R	Duration between the last purchase date of a customer and current time	The present time was assumed to be 31.10.2003.
F	The number of transactions throughout a customer's lifecycle	The total number of orders given by a customer was taken as a single value.
M	The revenue that is gained from a customer during lifecycle	The revenues of customers were determined as their monetary values.

The formula of RFM equals to F+M-R for calculating lifetime value of each customer (Liu and Shih, 2005b). Gelbrich and Wüschmann's Model (GWM) is in the form of flow money in between customer and enterprise.

$$CLV = \sum_{i=1}^n \frac{Ri - Ki}{(1+r)^i} \quad (\text{Equation 1})$$

The variables, operationalized in GWM formula is given below.

Table 2. Operationalization of the Variables for GWM

Variable	Explanation	Operationalization
n	Expected life of a customer	$n = \frac{1}{1-r}$ (Reicheld, 1996) value depends on the retention rate of customer.
R _i	Total revenue of customer in period <i>i</i>	The revenues of customers were assigned as their monetary values.
K _i	Total cost of customer in period <i>i</i>	<i>Distribution Cost:</i> Cost for each customer was assumed to be variable and it changes for each purchase, which can be formulated as followings: For each purchase if the number of products is between 1-50 then the cost is £12; 50-150 then the cost is £10; 150-300 then the cost is £6; 300-600 then the cost is £2; 600 and more then no charge
r	Discount rate (annual)	Assumed to be 30%.

4. Empirical Results

4.1. Lifetime Value Assessment and Segmentation

For the purposes of this essay, the procedure applied in this section contains some specific steps. At the beginning, lifetime value assessments or calculations of all customers were carried out and then the corresponding segments based on these values were generated. Regarding RFM model, labelling process for all customers was carried out using the operationalization given in Table 3 according to their R, F, and M values that were calculated separately for each of them. To be more accurate, each individual value for a customer was compared with the corresponding average value of all customers. If R (F, M) value of a customer was higher than the average R (F, M) values of all customers this particular customer was labelled as RH (FH, MH), while the R (F, M) value lower than the average R (F, M) was labelled as RL (FL, ML); where the second letters in the labels indicate the status of being high and low, respectively. In this way, with the aim of developing customer segments, eight different R-F-M combinations were generated. Subsequently, based on their R, F and M status, these combinations were classified into four groups. Table 4 gives information about four obtained segments and their descriptions together with number of customers in each dataset and the corresponding R-F-M combinations.

Table 3. Customer Segments and Descriptions

Segment	Description of the Segment	Number of Customers	Percent of Customers (%)
1	High Value Customers	220	7.30
2	Moderate-to-High Value Customers	1357	44.98
3	Low-to-Moderate Value Customers	1254	41.56
4	Low Value Customers	186	6.17

The other customer lifetime value model, GWM, lifetime value of each customer was calculated using Equation 1 provided in Table 2. Following this, in accordance with the corresponding calculated values, the consumers were sorted in a descending order for each model. To achieve an equivalent comparison base, in RFM and GWM models, the total numbers of segments were set equal to the segment structure generated by RFM model. Therefore, the first 220 customers in the ranking were described as “high value customers”, the followed 1357 of them as “moderate-to-high value customers”, the next 1254 of them as “low-to-moderate value customers” and the remaining 186 customers as “low value customers”.

4.2. Results of the Comparison

4.2.1 Separate Assessment of the Segmentation Results for Each Model.

Four different customer segments were obtained for two models. In order to ensure that the segments generated for each model can be identified according to the corresponding segmentation bases that were used during the segmentation process, ANOVA tests were performed at 0.05 level of significance for each segmentation structure and results were obtained as given in Table 4. From the figures it is apparent that the levels of significance for all corresponding variables of each model were found to be less than 0.05. For this reason, it can be said that for the related segments the average values of these variables were statistically different from each other. In other words, the segments obtained by the models are differentiable.

Table 4. Average CLV Values and Result of ANOVA tests for Each Model

Model	Segment 1	Segment 2	Segment 3	Segment 4	F	Sig
RFM	0,56	0,15	-0,15	-0,49	4970,71	0.00
GWM	334,867	141,91	57,67	20,82	720,65	0.00

4.2.2. Verification of the Differences between Segmentation Structures of Each Model.

Ensuring that the segmentation structure of each model is different from the other, the difference was set forth through calculating the similarity of the segmentation results. Cohen’s Kappa index was used to measure the agreement between the segmentation structures obtained. An index value converges to “0” indicates that the agreement between segmentation results is low, while a value close to “1” designates high level of agreement. However, any value between 0 and 1 can represent a certain level of agreement with a degree of randomness (Landis and Koch, 1977). The results of calculations demonstrated that the similarity percentage GWM and RFM were found to be 34%, respectively. (Table 5) It can be clearly seen that the segments generated by RFM and the segments obtained through GWM include different customers at a substantial amount. In another word, there is an observable defined pattern in the results of GWM compared to RFM model in terms of customers groupings. Therefore, it is

possible to distinguish or discern the segment structures of each model. Such differences would provide a basis for further comparison of the models.

Table 5. Cohen's Kappa index for Each Model

Model	RFM	GWM
RFM	1.00	0.34
GWM	0.34	1.00

4.2.3. Comparison of the Models from Segmentation Perspective.

The main objective of this research is to make a comparison of different lifetime value models at segment level for the purpose of discovering which one is superior to the others. The comparison was performed based on 'average revenues' of the segments using four control variables, namely, *value per visit* (average monetary value per visit/shopping), *unique product variety per visit* (number of unique products bought per visit/shopping), *quantity per visit* (total number of products bought per visit/shopping), and *unique product variety per quantity* (number of unique products bought over total number of products). Table 6, 7, 8, and 9 provide that information for each individual customer segment of the comparative models.

Table 6. The Calculation of Average Revenues of Customer Segments for Value per Visit

Segment Number	RFM	GWM	t	Sig.
1	62,08	137,44	-11,60	0.00
2	65,46	75,77	-5,21	0,00
3	55,60	36,04	13,97	0,00
4	47,85	15,34	10,93	0,00

Table 6 illustrate the results of calculations of average revenues and independent sample t-Test analysis of customer segments for control the variable of value per visit. By considering segment 1, it can be seen that the significance level is 0.00, which is below 0.05 and therefore there is a statistically significant difference in the mean value per visit between different segmentation structures generated by two different models. This conclusion is also valid for the results of segment 2, 3 and 4.

Table 7. The Calculation of Average Revenues of Customer Segments for Unique Product Variety per Visit

Segment Number	RFM	GWM	t	Sig.
1	10,67	24,54	-11,51	0.00
2	12,11	13,54	-4,23	0.00
3	10,20	7,31	12,34	0.00
4	8,75	3,71	11,03	0.00

Considering the control variable of unique product variety per visit, the independent sample t-Test results and average revenues by models were shown in Tables 7. All significance levels of segments are less than 0.05, so this means different models have different customer structure.

Table 8. The Calculation of Average Revenues of Customer Segments for Quantity per Visit

Segment Number	RFM	GWM	t	Sig.
1	13,42	28,68	-10,66	0.00
2	14,68	16,45	-4,17	0.00
3	12,37	8,70	12,69	0.00
4	10,70	4,42	10,69	0,00

Table 8 provides information on comparisons of the models from segmentation perspective for quantity per visit. The results are very similar to the ones that were presented in the previous tables. Also, the levels of significance for all corresponding variables of each group were found to be less than 0.05, therefore it can be inferred that the average values of these variables for the associated models were statistically different from each other.

Table 9. The Calculation of Average Revenues of Customer Segments for Unique Product Variety per Quantity

Segment Number	RFM	GWM	t	Sig.
1	26,44	30,08	-1,07	0,48
2	26,89	3,05	-3,65	0.00
3	31,85	28,87	3,02	0.01
4	30,57	23,30	4,22	0.00

According to the independent sample t-Test analysis, for the case of unique product variety per quantity (Tables 9), the differences between models calculated for segment 1 is not significant due to its *P* value (*P* values greater than 0.05 are insignificant); therefore there is insufficient evidence to claim that some of the means may be different from each other. In the other cases, all the differences between segments are meaningful.

The evidence from these results suggest that analyzing these segments cannot help reveal the differences between the comparative models. However, should one scrutinizes if there is a difference between the models based on Segment 1, s/he would figure out that the average revenues pertaining to valuable segment for GWM yields higher gain compared to the corresponding results of RFM model. The same results are also valid for calculations on Segment 2. On the contrary, when looking at the difference at Segment 3 GWM's average revenues seem to be lower in comparison with the associated results of their peers. General evaluations of differences lead us to the conclusion that the segmentation structures established by GWM were found to be more effective compared

to RFM model, since the GWM seem to be more capable of enabling the assignment of the most valuable customers into the same segment. This means that GWM has the ability to facilitate performing attraction of lucrative customers in one group and classifying the new customers in a lower value segment in a better way.

In order to understand better the segmentation structure of these two different models, two different bar charts were created and presented below (Figure 1, 2 and 3). Bar charts illustrate customer values of the segments generated by the models. The horizontal axes show the models while the vertical axes give information about some value-related indicators at shopper (customer) level, namely value (monetary) per shopper, unique product variety per shopper, quantity per shopper. Segments were represented via greyscale colors. These assessments were made through selection of 3017 customers' transactions from the available data. When the value per shopper is considered, GWM gave the highest value for Segment 1 in Figure 1. According to Figure 2, GWM yielded the highest value for Segment 1 based on unique product variety per. Additionally, when Figure 3 is analyzed, it can be seen that the models provided the same results similar to Figure 1 and Figure 2 in terms of quantity per shopper. In consideration of all the aforementioned discussions, it can be said that the use of GWM to measure customer value provide betters results and RFM values are not so far away to these results.

General evaluations of differences lead us to the conclusion that the segmentation structures established by GWM were found to be more effective compared to RFM model, since the GWM seems to be more capable of enabling the assignment of the most valuable customers into the same segment. This means that GWM has the ability to facilitate performing attraction of lucrative customers in one group and classifying the new customers in a lower value segment in a better way.

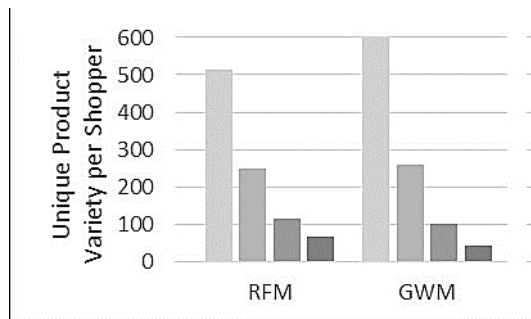


Figure 1. Illustration for “Customer Monetary Values per Shopper” of the Segments Generated by the Models

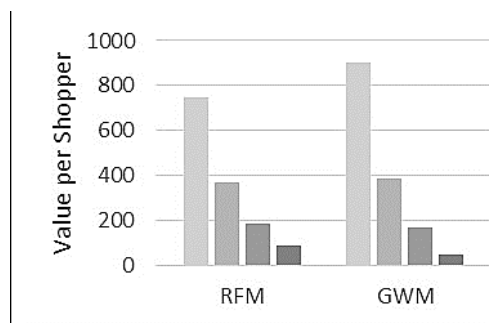


Figure 2. Illustration for “Unique Product Variety per Shopper” of the Segments Generated by the Models

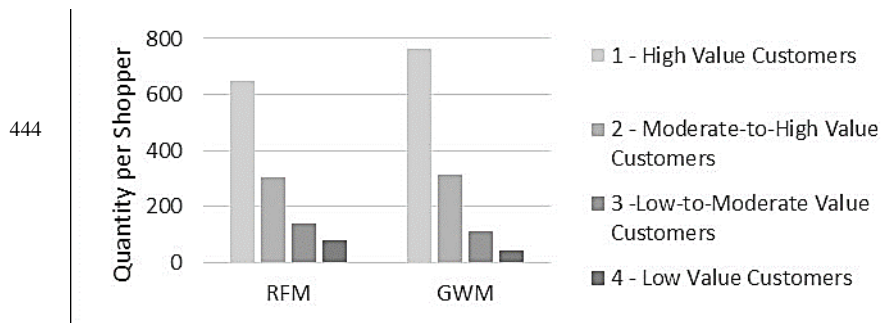


Figure 3. Illustration for “Quantity per Shopper” of the Segments Generated by the Models

5. Conclusion

Discovering differences between customers and specifying profitability of each customer have been one of the most important challenges in marketing. Firms can utilize CLV models in order to determine the characteristics of their customers. Moreover, through the means of customer segmentation, which could be carried out based on these value-based characteristics, organizations are able to develop appropriate strategies for supporting their decision making processes in customer relationship management context. This has become rather easy considering the availability of organized customer data and the approaches like data mining that can facilitate extraction of critical customer knowledge. Although the use of customer lifetime value for segmenting customers or formulating strategies tailored to them can be found in related literature, there has been a lack of comprehensive studies pertaining to analyzing different models and figuring out which model is superior to the others within data mining context. This study proposed a comparison to assess two different customer life time value models within data mining and from segmentation perspective by using value-related attributes as well as certain product-usage related control variables. In this context, at first, different CLV models were reviewed and two models that need the same set of variables were chosen for comparative assessment. One of these models is a past customer behavior model (RFM model), while the other model is future-past behavior model, Gelbrich and Wünschmann Model. Subsequently, the models were evaluated using the same data set based on the segmentation structure that they established. Comparisons were carried out based on ‘average revenues’ of the segments using four control variables via independent sample t-Test analyses. The results of the study demonstrated that GVM yielded better performance for all control variables and the segmentations obtained via this model could be seen more effective compared to RFM model.

In conclusion, the usage of CLV models and data mining techniques together gives a tremendous capability to the firms in recognizing high value customer groups. From this standpoint, this study provides two benefits to the current body of the literature as well as to the marketing practice. First, the article enhances academic understanding of existing CLV models from a taxonomic perspective. Second, the usage lifetime value and segmentation concepts within data mining context can provide a grasp of practical implementation in customer analytics area. In fact, comparison of the segmentation structures of two lifetime value models using four different control variables can facilitate a better comprehension from an empirical practice point of view. Nevertheless, a number of limitations of this study and areas for future research could also be mentioned. One limitation is that only a specific database was used to assess these models. It is far better that more analyses could have been performed on different

datasets for different types of sectors. In addition, another important point is that only two customer lifetime value models were utilized for comparisons since these models need the same set of variables. Other lifetime value models could have also been taken into account should it is possible to find common features for comparative assessment. Last but not least, some assumptions had to be kept in mind due to lack of specific consumer-related data/information. Making these assumptions more relaxed and building the research framework on obtaining data sets that could be more consistent with real conditions may ensure more robust results for future research.

References

- Aeron, H., Bhaskar, T., Sundararajan, R., Kumar, A., Moorthy, J. (2008). A metric for customer lifetime value of credit card customers. *Journal of Database Marketing and Customer Strategy Management*, 15(3), 153-168.
- Albadvi, A., Shahbazi, M. (2010). Integrating rating-based collaborative filtering with customer lifetime value: New product recommendation technique. *Intelligent Data Analysis*, 14(1), 143-155.
- Alvandi, M., Fazli, S., Abdoli, F. S. (2012). K-Mean clustering method for analysis customer lifetime value with LRFM relationship model in banking services. *International Research Journal of Applied and Basic Sciences*, 3(11), 2294-2302.
- Azadnia, A. H., Ghadimi, P., Aghdam, M. M. (2012). A Hybrid Model of Data Mining and MCDM Methods for Estimating Customer Lifetime Value. In *Proceedings of the 41st International Conference on Computers and Industrial Engineering*, Los Angeles, CA, USA (Vol. 2325, p. 8085).
- Bauer, H. H., Hammerschmidt, M., Braehler, M. (2003). The customer lifetime value concept and its contribution to corporate valuation. *Yearbook of Marketing and Consumer Research*, 1(1), 49-67.
- Bejou, D., Keiningham, T. L., Aksoy, L. (2006). *Customer lifetime value: Reshaping the way we manage to maximize profits*. Routledge.
- Benoit, D.F., Poel, D.V. (2009). Benefits of quantile regression for the analysis of customer lifetime value in a contractual setting: an application in financial services. *Expert Systems With Applications*, 36,10475-10484.
- Berger, P. D., Nasr, N. I. (1998). Customer lifetime value: Marketing models and applications. *Journal of interactive marketing*, 12(1), 17-30.
- Blattberg, R. C., Deighton, J. (1996). Manage marketing by the customer equity test. *Harvard business review*, 74(4), 136.
- Bruhn, M. (2003) *Relationship Marketing*. London, Pearson Education Limited,
- Chang, H. C., Tsai, H. P. (2011). Group RFM analysis as a novel framework to discover better customer consumption behavior. *Expert Systems with Applications*, 38(12), 14499-14513.
- Chang, H. H., Tsay, S. F. (2004). Integrating of SOM and K-mean in data mining clustering: An empirical study of CRM and profitability evaluation.
- Chen, C. W., Yang, C., Lin, C. S. (2009). A study of discovering customer value for CRM: Integrating customer lifetime value analysis and data mining techniques. *Information and Management Science*, 2(1), 14-30.
- Chen, Y. L., Kuo, M. H., Wu, S. Y., Tang, K. (2009). Discovering recency, frequency, and monetary (RFM) sequential patterns from customers' purchasing data. *Electronic Commerce Research and Applications*, 8(5), 241-251.
- Chen, Z. Y., Fan, Z. P. (2013). Dynamic customer lifetime value prediction using longitudinal data: An improved multiple kernel SVR approach. *Knowledge-Based Systems*, 43, 123-134.

- Cheng, C. H., Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. *Expert systems with applications*, 36(3), 4176-4184.
- Chuang, H. M., Shen, C. C. (2008). A study on the applications of data mining techniques to enhance customer lifetime value—based on the department store industry. In *Machine Learning and Cybernetics, 2008 International Conference on* (Vol. 1, pp. 168-173). IEEE.
- Collings, D., Baxter, N. (2005). Valuing customers. *BT Technology Journal*, 23(3), 24-29.
- Cuadros, A. J., Domínguez, V. E. (2014). Customer segmentation model based on value generation for marketing strategies formulation. *Estudios Gerenciales*, 30(130), 25-30.
- Drèze, X., Bonfrer, A. (2009). Moving from customer lifetime value to customer equity. *QME*, 7(3), 289-320.
- Dwyer, F. R. (1997). Customer lifetime valuation to support marketing decision making. *Journal of interactive marketing*, 11(4), 6-13.
- Ekinci, Y., Ülengin, F., Uray, N., Ülengin, B. (2014). Analysis of customer lifetime value and marketing expenditure decisions through a Markovian-based model. *European Journal of Operational Research*, 237(1), 278-288.
- Fader, P., Hardie, B., Berger, P. D. (2004). Customer-base analysis with discrete-time transaction data. Unpublished working paper. <http://dx.doi.org/10.2139/ssrn.596801>
- Fader, P. S., Hardie, B. G., Lee, K. L. (2005). RFM and CLV: Using iso-value curves for customer base analysis. *Journal of Marketing Research*, 42(4), 415-430.
- Freeman, M. (1999). The 2 customer lifecycles. *Intelligent Enterprise*, 2(16), 9.
- Gelbrich, K., Wünschmann, S., (2007). *Mehrdimensionaler Kundenwert als Entscheidungskriterium für die Akquisition von Kunden: Dargestellt am Beispiel der Automobilindustrie. Kundenwert: Grundlagen-Innovative Konzepte-Praktische Umsetzungen* pp 583-606 Gabler Verlag
- Glady, N., Baesens, B., Croux, C. (2009). A modified Pareto/NBD approach for predicting customer lifetime value. *Expert Systems with Applications*, 36(2), 2062-2071.
- Glady, N., Lemmens, A., Croux, C. (2015). Unveiling the relationship between the transaction timing, spending and dropout behavior of customers. *International Journal of Research in Marketing*, 32(1), 78-93
- Gloy, B. A., Akridge, J. T., Preckel, P. V. (1997). Customer lifetime value: An application in the rural petroleum market. *Agribusiness*, 13(3), 335-347.
- Golmah, V., Mirhashemi, G. (2012). Implementing A Data Mining Solution To Customer Segmentation For Decayable Products—A Case Study For A Textile Firm. *International Journal of Database Theory and Application*, 5(3), 73-90.
- Guo, Y., Wang, H., Liu, W. (2013). Improved pareto/nbd model and its applications in customer segmentation based on personal information combination. *International Journal of Database Theory and Application*, 6(5), 175-186.
- Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Sriram, N.R.S. (2006). Modeling customer lifetime value. *Journal of service research*, 9(2), 139-155.
- Gupta, S., Lehmann, D. R. (2003). Customers as assets. *Journal of Interactive Marketing*, 17(1), 9-24.
- Gupta, S., Lehmann, D. R., Stuart, J. A. (2004). Valuing customers. *Journal of marketing research*, 41(1), 7-18.
- Haenlein, M., Kaplan, A. M., Beeser, A. J. (2007). A model to determine customer lifetime value in a retail banking context. *European Management Journal*, 25(3), 221-234.
- Hoekstra, J. C., Huizingh, E. K. (1999). The lifetime value concept in customer-based marketing. *Journal of Market-Focused Management*, 3(3), 257-274.
- Hosseini, S. M. S., Maleki, A., Gholamian, M. R. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems with Applications*, 37(7), 5259-5264.

- Hosseni, M. B., Tarokh, M. J. (2011). Customer segmentation using CLV elements. *Journal of Service Science and Management*, 4(03), 284.
- Hu, Y. H., Huang, T. C. K., Kao, Y. H. (2013). Knowledge discovery of weighted RFM sequential patterns from customer sequence databases. *Journal of Systems and Software*, 86(3), 779-788.
- Hwang, H., Jung, T., Suh, E. (2004). An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry. *Expert systems with applications*, 26(2), 181-188.
- Jain, D., Singh, S. S. (2002). Customer lifetime value research in marketing: A review and future directions. *Journal of interactive marketing*, 16(2), 34-46.
- Jonker, J. J., Piersma, N., Van den Poel, D. (2004). Joint optimization of customer segmentation and marketing policy to maximize long-term profitability. *Expert Systems with Applications*, 27(2), 159-168.
- Khajvand, M., Tarokh, M. J. (2011). Analyzing customer segmentation based on customer value components (Case Study: A Private Bank)(Technical note).
- Khajvand, M., Zolfaghar, K., Ashoori, S., Alizadeh, S. (2011). Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study. *Procedia Computer Science*, 3, 57-63.
- Kim, B. D., Kim, S. O. (1999). Measuring upselling potential of life insurance customers: Application of a stochastic frontier model. *Journal of Interactive marketing*, 13(4), 2-9.
- Kim, S.Y., Jung, T.S., Suh, E.H. Hwang, H.S. (2006). Customer segmentation and strategy development based on CLV. *Expert Systems With Application*, 31, 101-107.
- Kim, W.G., Cha, Y. (2002). Antecedents and consequences of relationship quality in hotel industry. *International Journal of Hospitality Management*, 21(4), 321-338.
- Kumar, V. (2005). Uses, misuses, and future advance. In R. Grover M. Vriens,(Ed), *The handbook of market research* (pp. 602-628) California: Sage Publication
- Kumar, V., Shah, D. (2009). Expanding the role of marketing: from customer equity to market capitalization. *Journal of Marketing*, 73(6), 119-136.
- Kumar, V., Venkatesan, R., Bohling, T. Beckmann, D. (2008). The power of CLV: managing customer lifetime value at IBM. *Marketing Science*, 27, 585-599
- Lemon, K.N. Mark, T. (2006). Customer lifetime value as the basis of customer segmentation. *Journal Of Relationship Marketing*, 5, 55-69.
- Landis, J.R., Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174
- Lin, C. C., Shih, D. H. (2011). Data Mining Techniques to Enhance Customer Lifetime Value. In *Advanced Materials Research* (Vol. 225, pp. 3-7). Trans Tech Publications.
- Lin, S. Y., Wei, J. T., Weng, C. C., Wu, H. H. (2011). A case study of using classification and regression tree and LRFM model in a pediatric dental clinic. *International Proceedings of Economic Development and Research—Innovation, Management and Service*, 14, 131-135.
- Liu, D. R., Shih, Y. Y. (2005a). Hybrid approaches to product recommendation based on customer lifetime value and purchase preferences. *Journal of Systems and Software*, 77(2), 181-191.
- Liu, D. R., Shih, Y. Y. (2005b). Integrating AHP and data mining for product recommendation based on customer lifetime value. *Information Management*, 42(3), 387-400.
- Malthouse, E. C., Blattberg, R. C. (2005). Can we predict customer lifetime value?. *Journal of interactive marketing*, 19(1), 2-16.
- McDonald, M. A. (1996). Service quality and customer lifetime value in professional sport franchises.
- Parvaneh, A., Abbasimehr, H., Tarokh, M.J. (2012). Data mining application in retailer segmentation based on LRFM variables: case study. In *2nd World Conference on Information Technology (WCIT-2011)*. Antalya, Turkey,128-133

- Pfeifer, P. E., Carraway, R. L. (2000). Modeling customer relationships as Markov chains. *Journal of interactive marketing*, 14(2), 43-55.
- Ramakrishnan, R. (2006). Customer Lifetime Value. In: National Seminar On Changing Scenario Of Consumerism. Tiruchirapalli, India 07th-08th January 2006. Bharathidasan University: Department Of Commerce.
- Reinartz, W. J., Kumar, V. (2000). On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing. *Journal of marketing*, 64(4), 17-35.
- Rosset, S., Neumann, E., Eick, U., Vatnik, N. (2003). Customer lifetime value models for decision support. *Data mining and knowledge discovery*, 7(3), 321-339.
- Rust, R. T., Lemon, K.N., Zeithaml V.A., (2004), "Return on Marketing: Using Customer Equity to Focus Marketing Strategy," *Journal of Marketing*, 68 (1), 109-27.
- Venkatesan, R., Kumar, V. (2004). A customer lifetime value framework for customer selection and resource allocation strategy. *Journal of marketing*, 68(4), 106-125.
- Wiesel, T., Skiera, B., Villanueva, J. (2008). Customer equity: an integral part of financial reporting. *Journal of Marketing*, 72(2), 1-14.
- Wu, H. H., Lin, S. Y., Liu, C. W. (2014). Analyzing patients' values by applying cluster analysis and LRFM model in a pediatric dental clinic in Taiwan. *The Scientific World Journal*, 2014.
- Wu, S. I., Li, P. C. (2011). The relationships between CRM, RQ, and CLV based on different hotel preferences. *International Journal of Hospitality Management*, 30(2), 262-271.

Received October 31, 2018, accepted December 13, 2018