Analyzing Customer Segmentation Based on Customer Value Components (Case Study: A Private Bank) (Technical note)

Mahboubeh Khajvand *[1] and Mohammad Jafar Tarokh [1]
[1] IT Group, Industrial Engineering Department, K.N. Toosi University of Technology, Tehran, Iran

(Received 6 December 2010, Accepted 12 June 2011)

Abstract

Studying about the customer segmentation and begetting customer ranking plan diverts more attention in recent years. In this regard, this study tries on providing a methodology for segmenting customers based on their value driver parameters which extracted from transaction data and then ranks customers with regard to their customer lifetime value (CLV) score. Discovering hidden pattern between customers ranking result the other data such as customer product ownership data and socio-demographic information is the other work which addressed in this paper. Achieving this, we used data mining techniques such as different classification and clustering approaches, and implemented them on real data from an Iranian private bank. Current study can provide good insights for marketing and CRM department of the organization in identifying different segments of customer for designing future strategic program.

Keywords: Customer lifetime value, Customer segmentation, Data mining, RFM analysis, Decision rule

Introduction

Nowadays, with increase in market competition, more and more companies do realize that their most priceless asset is the existing customers [1], so they give more attention to customer relationship management (CRM) than past. The main goal of a CRM system is to understand profitable customers, to create and sustain relations with them. Thus it is so important to segment customers based on their value and dedicate rank to them and establish different relation with different segments of customer and different ranks. Customer lifetime value (CLV) as a paradigm of analytical CRM is considered by researchers and companies in different industries. CLV is the present value of all future profits for firms generated from a customer [2]. Calculating CLV has had lots of applications and several authors have developed models for different applications such as performance measurement [3], targeting customers [4], marketing resources allocation [5,6], product offering [7, 8, 9], pricing [10], and customer segmentation [11, 12, 13].

This study prepared to segment customers of an Iranian bank and based on it, the bank can focus on niche segment to propose new product and services to them which is one of the determined decisions in marketing strategy.

We considered three factors: Recency (R), Frequency (F), and Monetary (M) to cluster customers, analyzing clusters in RFM aspect, calculating CLV value of different clusters. Then clusters with homogeneous CLV value incorporate and construct a segment and based on the CLV value of the segments, we dedicate rank to them. After that decision tree and decision rules classifiers are used to estimating the accuracy rate of the customer segmentation and present rules for different segments to provide a clear and perspicuous knowledge explanation about different segments. Then socio-demographic predictors (e.g. age, sex) and product ownership predictors (e.g. types of customer accounts) are employed as input

* Corresponding author: Tel: +98-21-84063362 Tel: +98-21-88674858 Email: m.khajvand1984@gmail.com

variables to extract rule set for segments. These rules explain which customer with which characteristics lead to which segment with which value. Then evaluate which of the predictors have positive, negative, or neutral effect on customer CLV rank. Based on these extracted knowledge the bank could develop branch strategies like credit endowment or facility grant, marketing strategies, CRM strategies or even organization strategies for different segment of customer with different CLV rank.

The rest of this study is organized as follows. Section 2 outlines the background and reviews related work on customer lifetime value, CLV divisions and classifications, RFM analysis, and data mining definitions and techniques in customer segmentation application. Section 3 describes the research methodology and case study. Finally, section 4 draws conclusions and summarizing the contributions of this work.

1. Theoretical background

1.1. Customer lifetime value

Customer lifetime value (CLV) concept is going from customer relationship management (CRM) issue. Peppers believes that “the goal of CRM is to forge closer and deeper relationships with customers and to maximize the lifetime value of a customer to an organization” [14]. There has been an explosion of interest in the discipline and practice of CRM in the worlds of business and academic over the last decade especially in identifying and ranking of customers based on customer value drivers.

Based on the approach of estimating CLV, there are different definitions for this term. One of the earliest definition said, CLV is expected profits from customers except cost of customer management [15]. Pfeifer et al. defined CLV as the present value of the future cash flows attributed to the customer relationship [16] and finally Sublaban and Aranha [6] described CLV as estimated monetary value that the client will bring to the firm during the entire lifespan of his/her commercial relationship with the company, discounted to today value. In literature review, there are some classifications for CLV models. One of these divisions was proposed by Jain and Singh in [17] and the other by Gupta et al. in [18]. Jain and Singh determined that many models have been proposed in CLV literature dealing with all kinds of issues related to CLV. The following selection of models provides summaries of some key models addressing some major research opportunities in CLV research and applications. Based on the threefold stream of research related to CLV, they divided them into three corresponding categories [17]:

I. Models for calculation of CLV: This category includes models that are specifically formulated to calculate the CLV and/or extend this calculation to obtain optimal methods of resource allocation to optimize CLV.

II. Models of customer base analysis: Such models take into account the past purchase behavior of the entire customer base in order to come up with probabilities of purchase in the next time period.

III. Normative models of CLV: These models have been proposed and used mainly to understand the issues concerning CLV. Managers depend on many commonly held beliefs in making decisions regarding CLV.

Proposed paper works on normative model of Jain and Singh categories. The result of this research could be used by different department of the bank to make decision or plan strategy.

Gupta et al. described six modeling approaches in CLV issue [18]:

- RFM Models: Based on Recency, Frequency, and Monetary.
- Probability Models: Based on Pareto/NBD model and Markov chains.
- Econometric Models: Like probability model based on Pareto/NBD model and customer acquisition,

customer retention, and customer margin and expansion.

- Persistence Models: Based on modeling the behavior of its components, that is, acquisition, retention, and cross-selling.
- Computer Science Models: Based on theory (e.g., utility theory) and are easy to interpret. In contrast, the vast computer science literature in data mining, machine learning, and nonparametric statistics has generated.
- Diffusion/Growth Models: Based on customer equity (CE).

This study works on RFM model and uses computer science models' technique of Gupta's categories.

1.1.1. RFM analysis

One of the most powerful and simplest models to implement CRM may be the RFM model – Recency, Frequency, and Monetary value [19]. Bult and Wansbeek defined RFM as [20]: (1) R (Recency): the period since the last purchase; a lower value corresponds to a higher probability of the customer´s making a repeat purchase; (2) F (Frequency): number of purchases made within a certain period; higher frequency indicates greater loyalty; (3) M (Monetary): the money spent during a certain period; a higher value indicates that the company should focus more on that customer.

In recent researches, some authors proposed WRFM - Weighted RFM - instead of RFM. Depend on the importance of these parameters in their case they dedicated weights to R, F, and M. For example, Stone in [21] suggested placing the highest weighting on the Frequency, followed by the Recency, with the lowest weighting on the Monetary measure, in Liu and Shih [7] study, Recency is the most important parameter and Monetary is the less important parameter, but in Chuang and Shen study, Monetary has the most value and Recency had the least value [22].

To determine importance (weight) of RFM parameters, AHP method is exploited. The three main steps of this method are as follows ([]:

1. Perform pairwise comparisons with asking evaluators (decision makers or experts)
2. Assessing the consistency of pairwise judgments;
3. Employing eigen value computation to derive the weights of RFM variables.

Some researches try to develop RFM model and add some parameters to these three parameters. For example, Cheng Yeh et al. derived an augmented RFM model, called RFMTC model (Recency, Frequency, Monetary value, Time since first purchase, and Churn probability), using Bernoulli sequence in probability theory [19]. In this study, RFM parameters are employed for customer segmentation.

1.2. Data mining concept and methods

Simply stated, data mining is the process of automatically discovering useful information in large data repositories. Imielinski and Virmani described data mining as a pattern query in large data bases [23]. Data mining should have been more appropriately named knowledge mining from data. Han and Kumber believe that it is a step in the knowledge discovery process [24]. Knowledge discovery as a process is depicted in an iterative sequence of the following steps: Data cleaning; Data integration; Data selection; Data transformation; Data mining; Pattern evaluation; Knowledge presentation. The relation between the steps is shown in figure 1.

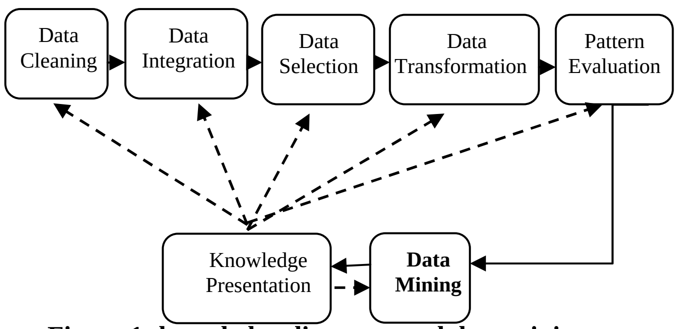


Figure 1: knowledge discovery and data mining process

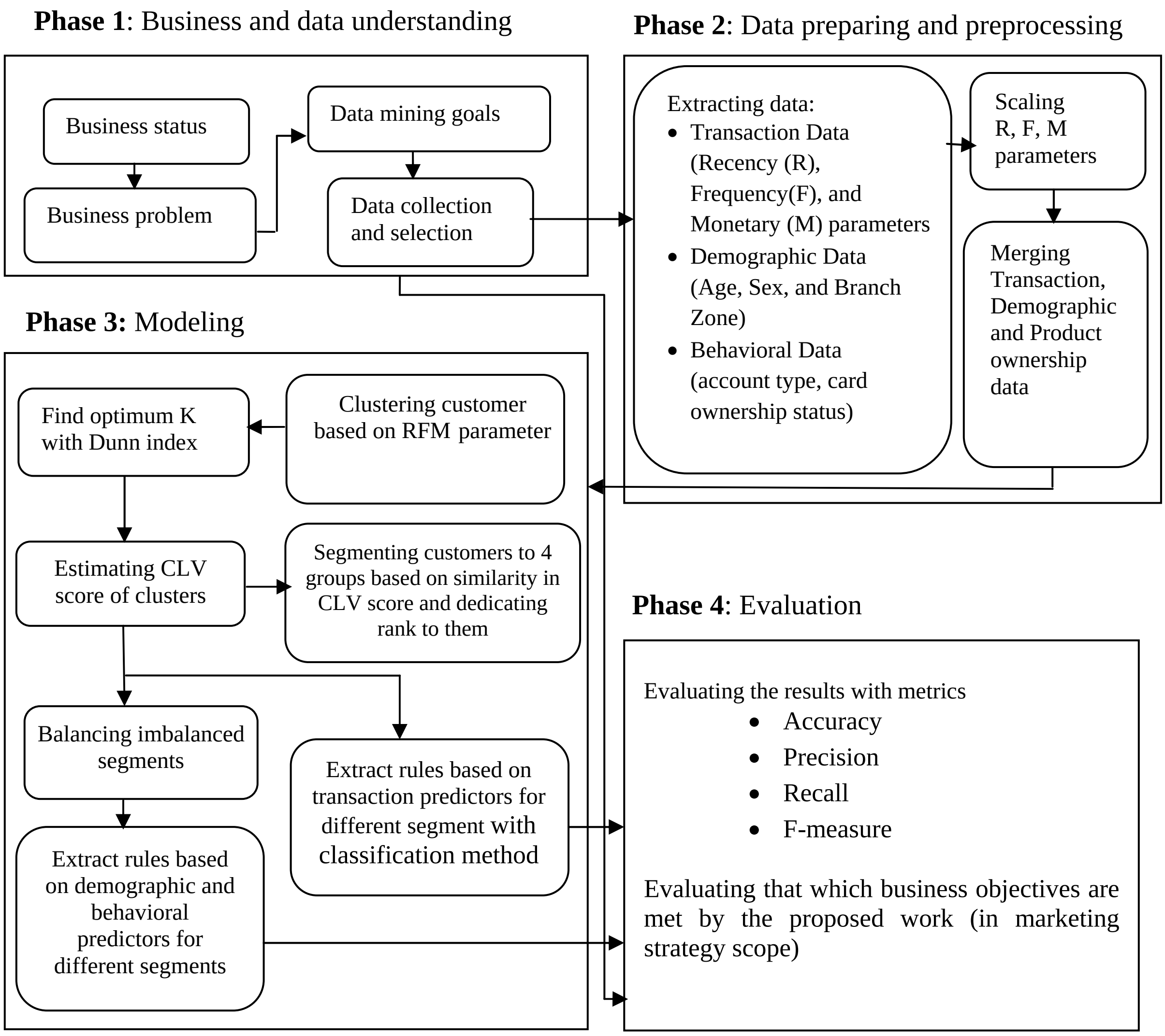


Figure 2: Structure of Research methodology

Data mining methods are two categories: descriptive and predictive. Descriptive method such as classification and predictive method such as clustering are used in current issue. Ngai et al. compare classification and clustering models as follow [25]: Classification aims at building a model to predict future customer behaviors through classifying database records into a number of predefined classes based on certain criteria, whereas clustering is the task of segmenting a heterogeneous population into a number of more homogenous clusters. It is different to classification in that clusters are unknown at the time the algorithm starts. In other words, there are no predefined clusters. Common tools used for classification are neural networks, decision trees and if then-else rules [12] and common tools for clustering include neural networks and discrimination analysis [26].

In CLV context, both of these models used in academic research, Baesens et al. used classification Bayesian network classifier in [27] and Drew et al. employed clustering neural network [28]. In this study, K-means algorithm as a clustering method and several decision tree algorithms as classification methods are used.

2. Methodology

2.1. Crisp methodology

There are different methodologies for implementing data mining projects but one of the powerful methods is CRISP (CRoss-Industry Standard Process for Data Mining) methodology [29]. As a process model, CRISP provides an overview of the data mining life cycle. CRISP uses six phases to describe the process from gathering business requirements to deploying the results [29]:

- Business Understanding: This phase typically involves gathering requirements and meeting with expert personnel to determine goals rather than working with data.
- Data understanding: The data understanding phase of CRISP involves taking a closer look at the data available for mining. This phase includes collecting initial data, describing data, exploring data, and verifying data quality.
- Data preparation: Data preparation is one of the most important and often time-consuming aspects of data mining projects and includes selecting data, cleaning data, constructing new data, and integrating data.
- Modeling: The data which was spent time preparing are ready to bring into data mining algorithms, and the results begin to shed some light on the business problem posed. Selecting modeling techniques, generating a test design, building the models, and assessing the model construct this phase.
- Evaluation: In this phase, evaluating the results, review process, and determining the next steps are done.
- Deployment: Deployment is the process of using the new insights to make improvements within the organization.

2.2. Research methodology

The CRISP methodology is flexible and can be customized easily. With regards to this methodology, a research methodology was proposed to ranking customers based on segmenting customer with customer value driver parameters. Figure 2 illustrate research methodology of this study.

Phase 1: Business and data understanding.

The case study concerns a private bank in Iran. It was established by a consortium of industrial, construction and investment companies, with the aim of providing flexible financial services to the burgeoning Iranian private sector. It has since grown to offer a full range of retail and commercial banking solutions, serving individual customers, small to medium sized businesses and large corporations and public institutions, both online and through our extensive national network of branches. Data mining department is a new department in this financial institute. This department tries to construct data warehouse from its huge data bases to replace analyzing on data based on OLTP (OnLine Transaction Process) system with OLAP (OnLine Analytical Process).

The most important goal of this organization is presentation of best services to customers. But the problem is that there is no taxonomy of customers based on their value in the bank. The main objective of current study is analysis on the customers' data to discover similar pattern between different segments of customer based on their equity and then assigning each segment of customer a rank based on their value, then extracting rule set for different segment lead to identifying high-value customers. The results of the study can then be used by the business to establish marketing and CRM strategy for different group of customers.

For this purpose, two years data of transactions, approximately five millions records for 50,000 customers, have been collected with simple stochastic sampling method. The data can be categorized roughly into three types, socio-demographic information, transaction information, and product ownership information. Table 1 explains some information about data.

Phase 2: Data preparing and preprocessing.

Extracting data, scaling R, F, M parameters, merging transaction, demographic, and product ownership data are sub phase of data preparing and preprocessing phase.

There are transaction data in detail which include type of transaction, date of transaction, monetary value of transaction (positive or negative), and type of deposit. For extracting R, F, and M parameters of each customer we aggregate records based on customer ID. In each aggregated record, there is a unique customer with her/his number of monetary transaction as F parameter, latest transaction date as R parameter and total amount of money in all of her/his deposit account at the end of the certain period (2 years) as M parameter. Then we scale each of these three parameters in five scales, very low, low, medium, high, and very high scale. For the scaling R parameter, first sort the data based on R attribute by ascendant order then partition the customers transaction dataset into 5 partitions. Customers in the first partition have lowest value of Recency and their Recency value named very low, second partition named low, third partition has medium scale, fourth is high scale and finally fifth partition named very high. For the scaling F and M parameters the same procedure must be done.

In the socio-demographic data, there is a field that called Branch zone, this attribute is derived from branch code. Tehran is divided into 22 zones (1 to 22) and has 95 branches, in the other hand, 111 branches are distributed to the other cities. These cities are divided into 6 zones (23 to 28) based on defined supervision zone in the bank.

In last step of this phase, we integrate data with merging socio-demographic, product ownership data and the result table of RFM aggregation. At the earliest, we have 50000 customers' data, but in the end of this phase we have a new aggregated data that include information (socio-demographic, product ownership, and aggregated transaction information) about 34000 customers.

Table 1: Data category and description

Data and information	Selected fields	Description
Socio-demographic information	Age	Is a range of 2 to 88
	Sex	0 for men and 1 for women
	Branch zone	There are 28 zones that 22 of them are in Tehran. It is derived from Branch code of each customer.
Transaction data	Recency	Latest transaction date during 2 years
	Frequency	Number of monetary transaction made within 2 years
	Monetary	Total amount of money in all of the customer's deposits in the end of the period
Product ownership data	Total product ownership	Total numbers of product
	Current account	1 for ownership 0 for non ownership
	Saving account	1 for ownership 0 for non ownership
	Short term time deposit accounts	1 for ownership 0 for non ownership
	Long term time deposit accounts	1 for ownership 0 for non ownership
	Debit card ownership	1 for card ownership 0 for non card ownership

.

Table 2: Scaling RFM parameters

Scaling	Scaling name	Recency	Frequency	Monetary
5 score	Very high (VH)	2010/1/6- 2010/1/10	[126,16968)	26,913,379 - 35,424,208,146
4 score	High (H)	2009/12/23- 2010/1/6	[56,126)	2,027,705 - 26,913,379
3 score	Medium (M)	2009/11/24- 2009/12/23	[27,56)	253,333- 2,027,705
2 score	Low (L)	2009/4/1- 2009/11/24	[8,27)	18.213 - 253,333
1 score	Very low (VL)	2008/1/9- 2009/4/1	[1,8)	0 - 18,213

Phase 3 and 4: Modeling and evaluation

As it is mentioned in first phase, the goal of using data mining tools in this study is ranking customers based on segmenting customers. In literature, customer segmentation is done in two ways: customer need-based and customer value-based segmentation [7, 8, 12, and 30]. In particular, customer needs-based segmentation must ultimately drive CRM programs even though a customer profitability segmentation analysis can serve as a good starting point for efficient program development [31].

For clustering customers, there are different algorithms which can be used. A variety of factors need to be considered when deciding which type of clustering technique to use [32]. One of the most common iterative algorithms is the K-means algorithm [33] which can be used for a wide variety of data types. It is also quite efficient, even though multiple runs are often performed. K-means has trouble clustering data that contains outliers. Outlier detection and removal can help significantly in such situations. Finally, K-means is restricted to data for which there is a notion of a center (centroid) [32]. At first step of modeling phase, we cluster customers based on customer value drivers (recency, frequency and monetary), with K-means algorithm. There are some indices for finding the optimum K (the number of clusters) that reviewed in [34]. We used Dunn index in this study [35]. The main goal of this measure is to maximize intercluster distances (distance between different clusters), whilst minimizing intracluster distances (distance between members of a cluster) [34]. For any partition $C = \{C_1, C_2, \ldots, C_k\}$, where C_i represents the i^{th} cluster of such partition, the Dunn indices, D, is defined as in equation (1):

$$D(C) = \frac{\min_{i,j=1,\ldots,k,\ i\neq j} \delta(C_i, C_j)}{\max_{i=1,\ldots,k} \Delta(C_i)} \tag{1}$$

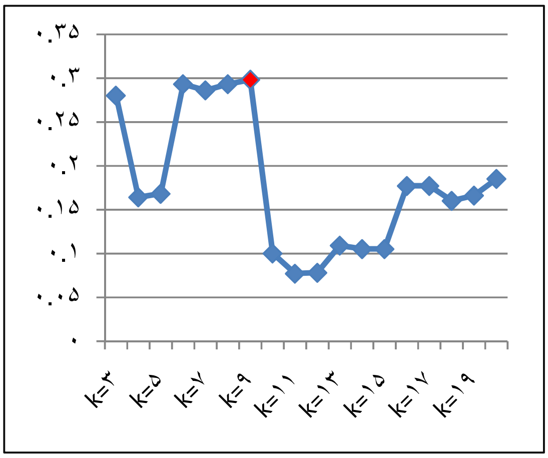
Where $\sigma(c_i, c_j)$ defines the distance between clusters c_i , c_j (intercluster distance), and $\Delta(c_i)$ represents the intracluster distance of cluster c_i or the size of the cluster c_j , and k is the total number of clusters. In this study:

$$\delta(C_i, C_j) = d(\bar{C}_i, \bar{C}_j) \tag{2}$$

$$\Delta(C_i) = \max_{x \in C_i}\{d(x, \bar{C}_i)\} \tag{3}$$

where $\bar{C}_i$ and $\bar{C}_j$ are the centroid of cluster C_i and C_j. Thus large Value of D corresponds to good clusters. Therefore, the number of cluster that maximizes D is taken as the optimal number of clusters, c. In this study optimum number of clusters (K) is 9 which calculated by Dunn index (Figure 3).

Figure 3: Optimum K with Dunn index



Second, based on centroid point of each cluster and information of table 2, R, F, and M score are dedicated to each cluster, as it is shown in table 3.

Table 3: Clustering customers with RFM parameters

cluster	No. of customer	R	F	M	RFM Analysis
c1	2965	2009/07/09	36	1,950,457	L - M - M
c2	5	2010/01/03	13689	934,639,633	H - VH - VH
c3	2758	2008/06/02	5	649,263	VL - VL - M
c4	391	2010/01/08	1873	472,360,068	VH - VH - VH
c5	4	2010/01/08	1443	31,854,247,414	VH - VH - VH
c6	3595	2008/10/17	15	641,384	VL - L - M
c7	20924	2009/12/29	120	98,573,791	H - VH - VH
c8	2663	2009/10/10	57	6,525,315	L - H - H
c9	2531	2009/03/06	23	1,887,606	VL - L - M

Table 4: RFM Score of clusters

cluster	% of customer	R Score	F score	M Score	RFM Score
c1	8.57	2	3	3	8
c2	0.01	4	5	5	14
c3	7.97	1	1	3	5
c4	1.13	5	5	5	15
c5	0.01	5	5	5	15
c6	6.81	1	2	3	6
c7	60.48	4	5	5	14
c8	7.70	2	4	4	10
c9	7.32	1	2	3	6

Table 5: Weights of RFM parameters

Parameter	R	F	M
Weight	0.087	0.345	0.653

Table 6: Calculating CLV with approach 2

Cluster	Normal R	Normal F	Normal M	CLV	Segmentation Ranking
c1	0.75	0.002080509	5.506E-05	0.066	3
c2	0.99	0.8067425	0.026384207	0.382	1
c3	0.20	0.000235752	1.83282E-05	0.017	4
c4	1.00	0.110355396	0.013334386	0.134	1
c5	1.00	0.084988507	0.899222568	0.703	1
c6	0.39	0.000801556	1.81058E-05	0.034	4
c7	0.98	0.00703719	0.002782667	0.090	2
c8	0.88	0.003300525	0.000184205	0.077	3
c9	0.58	0.001273059	5.32858E-05	0.051	4

RFM score of each cluster can be calculated with equation (4), as follow:

$$RFM\ Score_{ci} = R\ Score_{ci} + F\ Score_{ci} + M\ Score_{ci} \quad (4)$$

Where $R\ Score_{ci}$ refers to recency score of cluster ci, $F\ Score_{ci}$ is frequency score of cluster ci, and $M\ Score_{ci}$ is monetary score of cluster ci (table 4).

Cluster 2 and 7 with 20929 members which include 60 percents of all customers have 14 RFM score, other hand, C4 and C5 altogether have 395 members which include just one percent of all customers have 15 RFM score, as it is shown in figure 4.

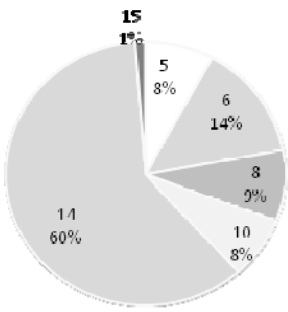


Figure 4: Diagram of customers' RFM score

After that, we calculate CLV score of each cluster based on weighted RFM analysis as follow in equation (5):

$$CLV_{ci} = NR_{ci} \times WR_{ci} + NF_{ci} \times WF_{ci} + NM_{ci} \times WM_{ci} \quad (5)$$

where NR_{ci} refers to normal recency of cluster ci, WR_{ci} is the weight of recency, NF_{ci} is normalized frequency, WF_{ci} is the weight of frequency, NM_{ci} is normalized monetary, and WM_{ci} is the weight of monetary. In this study, min-max normalization method is used for normalizing data. Min-max normalization performs a linear transformation on the original data [32]. Suppose that minA and maxA are the minimum and maximum values of an attribute, A. Then min-max normalization maps a value, v, of A to v' in the range of $[newmin_A, newmax_A]$ by computing:

$$v' = \frac{v - min_A}{max_A - min_A}(newmax_A - newmin_A) + newmin_A \quad (6)$$

In current study the new range is [0, 1]. According to the assessments obtained by the AHP method with regard to expert's insight of the bank, the relative weights of the RFM variables are mentioned in table 5.

In table 6 the normal R, F and M values for each cluster's centroid and CLV of them is shown.

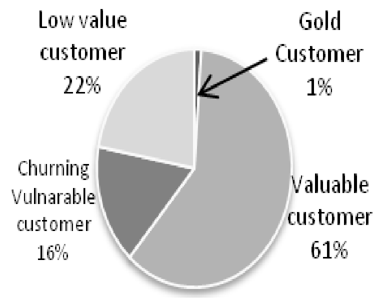


Figure5: Distribution of customers in segments

In general we have four segments of customers (as it is shown in last column of table 6). For analyzing customers' transaction behavior based on their value driver parameters, classifier algorithm can be used to extract rules. In fact, we extract rules for explanation of clear knowledge about different segments of customers with different ranks. In current study, several classifiers are used and PART (Projective Adaptive Resonance Theory) algorithm is chosen for extracting rules [36] because of its best metrics results (see table 8). PART uses separate-and-conquer and builds a partial C4.5 decision tree in each iteration and makes the best leaf into a rule. The comparison among decision tree classifiers and rule set classifiers is shown in table 8 and the rule set of PART decision rule technique is presented in list 1.

Metrics are used to guide the data mining algorithms and to evaluate the results of data mining. Recall and precision are two widely used metrics employed in applications where successful detection of one of the classes is considered more significant than detection of the other classes [32]. Precision and recall can be summarized into another metric known as the F-measure:

$$a = \frac{TP + TN}{TP + FP + TN + FN} \qquad Accuracy\ (7)$$

$$p = \frac{TP}{TP + FP} \qquad Precision (8)$$

$$r = \frac{TP}{TP + FN} \qquad Recall\ (9)$$

$$F = \frac{2 \times Recall \times Precision}{Recall + Precision} \qquad F - measure (10)$$

Where is:

- TP corresponds to the number of positive examples correctly predicted by the classification model
- FN corresponds to the number of positive examples wrongly predicted as negative by the classification model
- FP corresponds to the number of negative examples wrongly predicted as positive number of negative examples wrongly predicted as positive by the classification model
- TN corresponds to the number of negative examples correctly predicted by the classification model

Table 8: Comparison among classifiers

	Classifier	Accuracy (%)	Weighted Average precision	Weighted Average Recall	Weighted Average F-measure
Decision Tree	CHAID	95.11	0.948	0.953	0.950
	Decision Stump	82.56	0.721	0.826	0.76
Decision Rules	PART	99.95	1	1	1
	JRIP	99.94	0.999	0.999	0.999
	OneR	98.82	0.977	0.988	0.983

To discover the pattern between customers' value base segmentation (the last column of table 6) as *target field* and their socio-demographic and product ownership attributes as *predictors*, rule set algorithm can be used to extract rules. The algorithm which is exploited in current study is PART decision rule. Because of unbalancing problem among the four segments, all of the rules terminate in segment 2 with 60 percents of all customers. In a concept-learning problem, the data set is said to present a class *imbalance* if it contains many more examples of one class than the other. Such a situation poses challenges for typical classifiers such as decision tree induction systems or multi-layer perceptrons that are designed to optimize overall accuracy without taking in to account the relative distribution of each class. As a result, these classifiers tend to ignore small classes while concentrating on classifying the large ones accurately. There are some of the methods developed for handling the class imbalance problem [37]. Sampling-based methods and alternative metrics besides accuracy (i.e. recall and precision) introduced in this side. Data-based techniques for sampling include undersampling, oversampling, and a hybrid of both approaches. In the current study, we used hybrid one to reduce overfitting problem which is the result of oversampling method and omitting certain important records that is the result of undersampling method.

As it is mentioned prior, predictors fields include socio-demographic and product ownership attributes and the target field is the segment number. So, we run PART decision rule two times, first with product ownership data as predictors and second with socio-demographic data. Survey on results shows that socio-demographic data of customers are not good predictors of the segment of customers, because the accuracy of the rules is about 37 percents, that is very low. But product ownership data could be proper predictors for customer base segmentation. Before sampling, all of the rules terminated in segment 2 with 60 percents of all customers but after balancing data, we could extract a rule set for product ownership variables and the segmentation of the customers, which cover all segments (see list 2). The discovered knowledge is summarized in table 9. The content of table 9 show that Current account, the total number of product ownership and long term time deposit accounts have a positive relation with CLV score of customers. Also socio-demographic data such as age, sex and branch zone have no relation with customer value score.

2.2.1. Segmentations ranking analysis and proposed strategy

Segment 1: First segment includes the most valuable customers of the financial institute. This group is constructed from C2, C4, and C5 which have high CLV value. These customers are labeled “gold customers” and just one percent of customers belong to this segment. The bank must have special program and strategy for them. Private banking that is one of the three scopes of banking marketing strategies, must be consider for this segment. The bank should give consultant services to this group for managing their property.

Segment 2: Cluster 7 with 60 percents of customers’ contribution constructs this segment. This group contains customers who averagely refer to the bank one time in a week and the last transaction date of them is about ten days ago and average total monetary value of them is very high score of monetary scoring in retail banking scope. So this segment includes valuable customers of the bank and the institute must specify their retail banking strategy based on these customers. This means that new products should be developed based on this segment interests.

Segment 3: Segment 3 contains customers of C1 and C8 clusters, which includes 16 percents of customers. The members of this group like fourth group don’t have considerable recency but their frequency and monetary attribute have better score. Thus these customers have 3 states: they may incorporate into competitors or perhaps they don’t work with their account and finally they maybe just have saving accounts. The bank should suggest Current account to them to keep this group of customer, because of positive relation between this product and CLV score of customers.

Segment 4: Fourth segment includes C3, C6, and C9 that have 5 and 6 RFM Score. All of these clusters include customers with very low scale of recency that expose vulnerability of churning. These customers also have low or even very low scale of frequency which includes 22 percents of all customers, so they are disloyal customers. If the bank doesn’t want to absorb customers, it could be unconsidered to them and focused on the customers who belong to first and second segments.

Table 9: Relationship between product ownership, socio-Demographic and CLV Score

Predictor category	Explanatory variable	Relationship with CLV Score
Product ownership	Number of product ownership	Positive
	Current account	Positive
	Saving account	Neutral and Positive
	short term time	Neutral
	short term time deposit accounts	Neutral and Negative
	long term time deposit accounts	Positive
	Debit Card ownership	Neutral and Positive
Socio-Demographic	Age	Neutral
	Sex	Neutral
	Branch zone	Neutral

List 1: PART decision list for RFM

1. **Recency > 2009-11-26 AND Frequency > 855: Segment 1 (Rank=1)**
2. **Monetary > 300520712 AND Monetary <= 15120531861: Segment 2 (Rank=2)**
3. **Recency > 2009-11-19 AND Frequency <= 976 AND Monetary <= 5117099633: Segment 2 (Rank=2)**
4. **Recency <= 2009-11-26 AND Frequency <= 660 AND Recency <= 2009-11-19: Segment 3 (Rank=3)**
5. **Monetary <= 11229573246 AND Frequency <= 2547: Segment 3 (Rank=3)**

List 2: PART decision list

1. **Current account= 1 AND Short term time= 1: Segment 1**
2. **Long term time deposit accounts = 1 AND Card ownership = 0 AND Number of product ownership= 2: Segment 2**
3. **Short term time= 1 AND Short term time deposit accounts = 1 AND Card ownership = 1: Segment 2**
4. **Long term time deposit accounts = 1 AND Card ownership = 0 AND short term time= 1: Segment 2**
5. **Long term time deposit accounts = 1 AND Short term time= 0 AND Number of product ownership= 2: Segment 3**
6. **Long term time deposit accounts = 1 AND Short term time= 0 AND Number of product ownership = 1: Segment 4**
7. **Long term time deposit accounts = 0 AND Short term time= 0 AND Current account= 0: Segment 4**
8. **Current account= 0 AND Long term time deposit accounts = 0 AND Card ownership = 0 AND saving account= 0: Segment 4**
9. **Short term time= 1 AND Number of product ownership= 1: Segment 4**

3. Discussion and Conclusion

Not respecting to distinguish between customers lead to allocating recourses to all customers without considering returning on investigation. Thus it is so important and vital for firms to know different types of customers to make decision more profitability. This study provides new knowledge about ranking of a private bank customer based on their equity or in the other word CLV. Clustering customers into different groups not only leads to ranking customers based on customer equity parameters but also helps decision-makers identify market segments more clearly and thus develop more effective strategies based on transaction and product ownership data of customers. Customer lifetime value (CLV) has lots of application in marketing and CRM field. The marketing strategy could be summarized in 3 scopes in the bank: private banking, retail banking and corporate banking. In retail banking scope, the bank marketing department determined strategy plan which includes: evaluating of branches performance, ranking of bank customers, interviewing with gold customers, marketing the new products and services and etc. The result of this research could be used for all of the mentioned plans, because customer segmentation based on their value could be an appropriate measure in comparison of branches performance. In the other hand, ranking of customers based on their equity is one of the current research results, thus identifying the gold customer in retail banking scope is completely possible. Finally, for marketing the new products and services, marketing department should know which group of customer interest in which type of products and services, and so prediction of customer CLV rank based on customer products ownership that is the final result of the study, can provide a good insight for making decision in product and service design scope.

References:

1- Van den Poel, D. and Bart, L. (2004). “Customer attrition analysis for financial services using proportional hazard models.” *European Journal of Operational Research*, 157, PP. 196-217.

2- Gupta, S. and Lehmann, D.R. (2003). “Customers as assets.” *Journal of Interactive Marketing*, 17(1), PP. 9 – 24.

3- Rust, Lemon and Zeithaml. (2004),. “Return on marketing: using customer equity to focus marketing strategy.” *Journal of Marketing, 68*, 1, PP. 109-127.

4- Haenlein, M., Kaplan, A. M. and Schoder, D. (2006). “Valuing the real option of abandoning unprofitable cstorners when calculating customer lifetime value.” *Journal of Marketing, 70*, PP. 5-20.

5- Reinartz, W. J., Thomas, J. S. and Kumar, V. (2005). “Balancing acquisition and retention resources to maximize customer profitability.” *Journal of Marketing, 69*, PP. 63-79.

6- Sublaban, C.S.Y. and Aranha, F. (2009). “Estimating cell phone providers’ customer equity.” *Journal of Business Research*, 62, PP. 891-898.

7- Liu, D. R. and Shih, Y. Y. (2005a). “Hybrid approaches to product recommendation based on customer lifetime value and purchase preferences.” *The Journal of Systems and Software,77*, PP. 181-191.

8- Liu, D. R. and Shih, Y.Y. (2005b). “Integrating Ahp and data mining for product recommendation based on customer lifetime value.” *Information & Management , 42*, 14.

9- Shih, Y. Y. and Liu, D. R. (2007). “Product recommendation approaches: collaborative filtering via customer lifetime value and customer demands.” *Expert Systems with Applications*, 11.

10- Hidalgo, P., Manzur, E., Olavarrieta, S. and Farías, P. (2007). “Customer retention and price matching: The Afps case.” *Journal of Business Research, 61*, 6, PP. 691-696.

11- Rosset, S., Neumann, E., Eick, U., Vatnik, N. and Idan, Y. (2002. "Customer lifetime value modeling and its use for customer retention planning." *Paper presented at the SIGKDD '02*.

12- Kim, S.Y., Jung, T.S., Suh, E.H. and Hwang, H.S. (2006). "Customer segmentation and strategy development based on customer lifetime value: A case study." *Expert Systems with Applications*, 31, PP. 101–107.

13- Haenlien, M., Kaplan, A. M. and Beeser, A. J. (2007). "A model to determine customer lifetime value in a retail banking context." *European Management Journal*, 3, 5, PP. 221-234.

14- Peppers, D., Rogers, M. and Dorf, R. (1999). "Is your company ready for one-to-one marketing." *Harvard Business Review* 77, PP. 151–160.

15- Blattberg, R. C. and Deighton, J. (1996). "Manage marketing by the customer equity test." *Harvard business review*, 74, 4, PP. 136-144.

16- Pfeifer, P. E. and Bang, H. (2005). "Non-parametric estimation of mean customer lifetime value." *Journal of Interactive Marketing 19*, 4, PP. 48-66.

17- Jain, D. and Singh, S. S. (2002). "Customer lifetime value research in marketing: a review and future directions." *Journal of Interactive Marketing, 16*, 2, PP. 34-46.

18- Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V. and Lin, N. (2006). "Modeling customer life-time value." *Journal of Service Research*, 9, 2, PP. 139-155.

19- Cheng Yeh, Yang, K. J. and Ting, T. M. (2009). "Knowledge discovery on RFM model using Bernoulli sequence." *Expert Systems with Applications,* 36, PP. 5866-5871.

20- Bult, J. R. and Wansbeek, T. (1995). "Optimal selection for direct mail." *Marketing Science*, 14, PP. 378-395.

21- Stone, B. (1995). "Successful direct marketing methods." *NTC Business Books, Lincolnwood*.

22- Chuang, H. M. and Shen, C. C. (2008). "A study on the application of data mining techniques to enhance customer lifetime value—based on the department store industry." *The seventh international conference on machine learning and cybernetics*, PP. 168-173.

23- Imielinski, T. and Virmani, A. (1999). "MSQL – query language for data mining applications." data *mining and Knowledge Discovery Journal*.

24- Han, J. and Kamber, M. (2006). "Data Mining: Concepts and Techniques." CA*: Morgan Kaufmann, San Francisco*.

25- Ngai, E.W.T., Xiu, L. and Chau, D.C.K. (2009). "Application of data mining techniques in customer relationship management: A literature review and classification." *Expert Systems with Applications,* 36, PP. 2592–2602.i

26- Lee, J. H. and Park, S. C. (2005). "Intelligent profitable customers segmentation system based on business intelligence tools." *Expert Systems with Applications*, 29, PP. 145–152.

27- Baesens, B., Verstraeten, G., Dirk, V. D. P., Michael, E. P., Kenhove, V. K. and Vanthienen, J. (2004). "Bayesian network classifiers for identifying the slope of the customer-lifecycle of long-life customers," *European Journal of Operational Research*, 156, PP. 508–523.

28- Drew, J. H., Mani, D. R., Betz, A. L. and Datta, P. (2001). "Targeting customers with statistical and data-mining techniques." *Journal of Service Research,* 3, PP. 205–220.

29- Larose, D.T. (2006). "Data mining methods and models." *John Wiley & Sons, Inc., Hoboken, New Jersey*.

30- Hosseini, S.M.S., Maleki, A. and Gholamian, M.R. (2010). “Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty.” *Expert Systems with Applications*, 37, PP. 5259–5264.

31- Ravi, V. (2008). “Advances in banking technology and management: impacts of ICT and CRM.” *Information science reference, Hershey, New York, Yurchak Printing Inc*.

32- Tan, P.-N., Steinbach, M. and Kumar, V. (2005). Introduction to data mining. *pearson education*.

33- Brun, M., Johnson, C.D. and Ramos, K.S. (2005). “Clustering: revealing intrinsic dependencies in microarray data.” *Genomic Signal Processing and Statistics, EURASIP Book Series on Signal Processing and Communications, Hindawi Publishing Corporation*, PP. 129–162.

34- Brun, M., Sima, C., Huaa, J., Loweya, J., Carroll, B., Suha, E. and Doughertya, E. (2007). “Model-based evaluation of clustering validation measures.” *Pattern Recognition,* 40, PP. 807 – 824.

35- Dunn, J. (1974). “Well separated clusters and optimal fuzzy partitions.” *J.Cybernetics*, 4, PP. 95-104.

36- Cao, Y. and Wu, J. (2002). “Projective ART for clustering data sets in high dimensional spaces.” *Neural Networks*, 15, PP. 105-120.

37- Batista, G.E.A.P.A., Prati, R.C. and Monard, M.C. (2004). “A study of the behavior of several methods for balancing machine learning training data.” *SIGKDD Ex-plorations,* 6(1), PP. 20-29.