Customer Lifetime Value (CLV) Measurement Based on RFM Model

Babak Sohrabi[1], Amir Khanlari[2]

[1.] Assistant professor, School of Management, University of Tehran
[2.] PhD student of Marketing, School of Management, University of Tehran

Abstract
Nowadays companies increasingly derive revenue from the creation and sustenance of long-term relationships with their customers. In such an environment, marketing serves the purpose of maximizing customer lifetime value (CLV) and customer equity, which is the sum of the lifetime values of the company's customers. A frequently-encountered difficulty for companies wishing to measure customer profitability is that management accounting and reporting systems have tended to reflect product profitability rather than customer profitability. But in spite of these difficulties, Companies looking for methods to know how calculate their customers's CLV. In this paper, we used K-Mean clustering approach to determine customers's CLV and segment them based on recency, frequency and monetary (RFM) measures. We also used Discriminant analysis to approve clustering results. Data required applying this method gathered from one branch of an Iranian private bank which is established newly. Finally, in terms of this segmentation, we proposed customer retention strategies for treating with the bank customers.

Corresponding author: Email: bsohrabi @ ut.ac.ir

1. Introduction

Traditional accounting practices focus mainly on measuring tangible assets as a statutory requirement on the balance sheet. However, nowadays it is more usual for intangible assets such as brand, employee and customer relationships to be the critical and often dominant determinants of shareholder value [1]. A frequently-encountered difficulty for companies wishing to measure customer profitability is that management accounting and reporting systems have tended to reflect product profitability rather than customer profitability [2]. Meanwhile most companies have accounting systems that track costs based on functions (e.g., freight) rather than on a per customer basis [20].

As a result, medium-volume customers tend to be the most profitable. Unfortunately, standard accounting systems focus on periods instead of individual customers or customer groups [16]. To avoid such twists, customers need to be treated as a bundle of cost drivers. This is precisely the principle of Activity-Based Accounting (ABC) [11]. It implies that customers are the cause of activities and resources are employed to carry out activities to serve them. Costs are thus allocated on the basis of transactions. ABC therefore provides a fairly accurate means of measuring costs related to customer relationships.

For years, the challenge for businesses could largely be seen as putting in place the means of production to satisfy growing demand, and using marketing techniques to capture customers entering the market (e.g., [7, 18, 30]). Manufacturers of goods today, however, are competing in a very different environment, and transaction marketing (product, price, place, and promotion, the 4Ps) alone is believed to be insufficient [13, 36]. Instead, relationship marketing is proposed for building more unique relationships with customers and for adding more value to goods and services than what is possible through transaction marketing [17, 28]. Relationship marketing, then, is not only about the 4Ps but also long- term relationships, reflecting a transaction- relationship continuum [38]. Relationship marketing constitutes a major shift in marketing theory and practice. Rather than focusing on discrete transactions, it emphasizes the establishment, development and maintenance of long-term exchanges [29]. Such relationships are thought to be more profitable than short-term relationships as a result of exchange efficiencies. This is especially true of customer relationships [33].

However, since not all customers are financially attractive to the firm, it is crucial that their profitability be determined and that resources be allocated according to the customer lifetime value (CLV). There are several factors that account for the growing interest in this concept. First, there is an increasing pressure in companies to make marketing accountable. Second, financial metrics such as stock price and aggregate profit of the firm or a business unit do not solve the problem either. Although these measures are

useful, they have limited diagnostic capability. Third, improvements in information technology have made it easy for firms to collect enormous amount of customer transaction data. This allows firms to use data on revealed preferences rather than intentions [20].

Nonetheless, companies confront to CLV calculation problems yet. Although there is many models for this purpose, but most of them are theoretic, complex and not applicable. It is also important to point out that most modeling approaches ignore competition because of the lack of competitive data. Finally, how frequently we update CLV depends on the dynamics of a particular market. For example, in markets where margins and retention may change dramatically over a short period of time (e.g., due to competitive activity), it may be appropriate to re-estimate CLV more frequently. Therefore, companies need an approach to be used easily and its input data can be gathered and updated speedy. The proposed approach in this paper attempts to satisfy these expectations. This paper aims at suggesting a new CLV model and customer segmentation considering RFM model. We will also propose Customer retention strategies after segmenting customer base. For this purpose, the plan for the article is as follows. We first present CLV concept and its applications and what are its key drivers. Next, we present several modeling approaches that have been adopted to calculate CLV. This is followed by a detailed discussion of RFM as model used in this work. Then, we present clustering outputs and strategies suggested to treat with each customer cluster. We end the article with concluding remarks.

2. Customer Lifetime Value (CLV)

Customer value has been studied under the name of LTV, Customer Lifetime Value, Customer Equity, and Customer Profitability. The previous researches contain several definitions of CLV. The differences between the definitions are small [24]. Table 1 shows the definitions of CLV. Considering the definitions above, we define CLV as the sum of the revenues gained from company's customers over the lifetime of transactions after the deduction of the total cost of attracting, selling, and servicing customers, taking into account the time value of money.

In other words, CLV is generally defined as the present value of all future profits obtained from a customer over his or her life of relationship with a firm. CLV is similar to the discounted cash flow approach used in finance. However, there are two key differences. First, CLV is typically defined and estimated at an individual customer or segment level. This allows us to differentiate between customers who are more profitable than others rather than simply examining average profitability. Second, unlike finance, CLV explicitly incorporates the possibility that a customer may defect to competitors in the future [20].

   It is argued that customer relationships are viewed as investment decisions and customers as generators of revenue streams. Customer relationships also generate costs. Hence, in order to measure the customer lifetime value, all revenues and costs pertaining to a customer relationship must be assessed. It is then possible to calculate the current value of cash flow streams [4]. Though, accurately estimating the revenues and costs of a relationship remains a challenging task for a number of reasons:

Table 1: Definitions of CLV [20]

Definition	Reference
The present value of all future profits generated from a customer	[19]
The net profit or loss to the firm from a customer over the entire life of transactions of that customer with the firm	[4]
Expected profits from customers, exclusive of costs related to customer management	[6]
The total discounted net profit that a customer generates during her life on the house list	[5]
The net present value of the stream of contributions to profit that result from customer transactions and contacts with the company	[31]
The net present value of a future stream of contributions to overheads and profit expected from the customer	[25]
The net present value of all future contributions to overhead and profit	[34]
The net present value of all future contributions to profit and overhead expected from the customer	[12]

- Standard accounting does not allow for allocating costs to specific customer relationships.
- Only monetary benefits of customer relationships are taken into account.
- Revenues and costs vary over time.
- Cash flow streams are generated at different points in time and at different levels of risk.

   Consequently, the following requirements have to be fulfilled in order to measure customer profitability accurately [35]:
- An exact allocation of costs to customer relationships according to the resources employed;

- An estimation of all monetary and nonmonetary benefits created by the particular customer relationship;
- A consideration of cost and revenue changes over the estimated time span of a customer relationship;
- The discounting to present of future cash flows generated over the estimated time span of a customer relationship; and
- An estimation of the relationship risks.

There are a large number of models which attempt to calculate CLV. At the next section, some of most important models will be describe and then RFM model which seem to be more simply and robust will be present in more detail.

2.1. Models of CLV calculation

There are a lot of researches on calculating customer value. The basic concept of these researches, however, focused on Net Present Value (NPV) obtained from customers over the lifetime of transactions [3, 4, 19, 34]. Dwyer (1997) tried to calculate CLV through modeling the retention and migration behavior of customers. Focused on making decision of marketing invest, Hansotia and Rukstales (2002) suggested incremental value modeling using tree and regression based approach. Hoekstra and Huizingh (1999) also suggested a conceptual CLV model and categorized input data of the model into two types, source of interaction data and time frame. Most CLV models stem from the basic equation, although we have many other CLV calculation models having various realistic problems. The basic model form based upon the proposed definition is as follows

$$CLV = \sum_{i=1}^{n} \frac{(R_i - C_i)}{(1+d)^{i-0.5}}$$

Where i is the period of cash flow from customer transactions, R_i the revenue from the customer in period i, C_i the total cost of generating the revenue R_i in period i, and n is total number of periods of projected life of the customer under consideration. Therefore, the numerator is the net profit that has been obtained at each period while the denominator transforms the net profit value into the current value. The calculation model above is the most basic model that ignores the fluctuation of sales and costs. Expanding this basic model, many researchers including Berger and Nasr (1998) have proposed CLV calculation models, which reflect the fluctuation of sales and costs [6, 26]:

$$CLV = \sum_{i=0}^{n} \pi(t) \times \frac{1}{(1+d)^i}$$

Where $\pi(t)$ is the function of customer profits according to time t. formulating precise $\pi(t)$ is the most important factor in calculating CLV precisely. Colombo and Jiang (1999) developed a stochastic Recency Frequency Monetary model to rank customers in terms of their expected contribution. Pfeifer and Carraway (2000) proposed Markov chain models for modeling customer relationships.

The evaluations of customer value in previous studies have treated prediction method with regression models simply based on profits from customers to calculate the future value of customers. That is to say, considering the changing profit contribution obtained from customers in the past, the existing models calculate the future worth and then define the CLV of customers with the projected value of the future worth. Therefore, the CLV model above is not capable of considering potential values of customers, not available from the past profit contribution, which would be able to be the profits of companies.

Last, the models mentioned above do not considered the defection of customers. Although we have a customer who has very high value to our company, this information can conclude improper marketing strategies if we don't pay careful attention to the possibility of the customer defection. Hence, it is reasonable to consider the probability of individual customer's churn rather than to consider only the total decreasing rate of whole customers. Verhoef and Donkers (2001) used two dimensions, current value and potential value, to segment the customers of an insurance company.

2.2. RFM Model

To identify customer behavior, the well known method called recency, frequency and monetary (RFM) model is used to represent customer behavior characteristics [9, 23]. RFM models have been used in direct marketing for more than 30 years. Given the low response rates in this industry (typically 2% or less), these models were developed to target marketing programs (e.g., direct mail) at specific customers with the objective to improve response rates. Prior to these models, companies typically used demographic profiles of customers for targeting purposes. However, research strongly suggests that past purchases of consumers are better predictors of their future purchase behavior than demographics [20].

The basic assumption of using the RFM model is that future patterns of consumer trading resemble past and current patterns. The calculated RFM values are summarized to clarify customer behavior patterns. This study proposes using the following RFM variables [9]:

- Recency (R): the latest purchase amount.
- Frequency (F): the total number of purchases during a specific period.
- Monetary (M): monetary value spent during one specific period.

As mentioned, this approach models three dimensions of customer transactional data to classify customer behavior [39]. The first dimension is recency, which indicates the length of time since the start of a transaction. Meanwhile, the second dimension is Frequency, which indicates how frequently a customer purchases products during a particular period. Finally, monetary value measures the amount of money that customer spending during a period [27].

A large number of studies specifically in loyalty programs areas considered RFM. For instance, Jonker et al. (2004) demonstrated the use of RFM value in direct-mailing; they proposed an optimization strategy for customer segmentation and marketing employing stochastic dynamic programming. Also, Buckinx and Van den Poel (2005) and Fader et al. (2005) proved that RFM variables can predict accurately the CLV. They showed how RFM variables can be used to build a CLV model that overcomes many of its limitations. They also showed that RFM are sufficient statistics for their CLV model.

3. Research Methodology

We selected 214 RFM data from customer base of the given case. For data analysis, K- Means Cluster Analysis is used to cluster the Bank's customers. This procedure attempts to identify relatively homogeneous groups of cases based on selected characteristics, using an algorithm that can handle large numbers of cases. However, the algorithm requires you to specify the number of clusters. The parameter was set to 8, since eight $(2 \times 2 \times 2)$ possible combinations of inputs (RFM) can be obtained by assigning high or low, according to the average R (F, M) value of a cluster being less than or greater than the overall average R (F, M). The RFM values of customers were normalized as follows. The profit form, $x' = (x - x^S)/(x^L - x^S)$, was used to normalize the F (frequency) and M(monetary) values, since F and M positively influenced CLV or loyalty. The cost form, $x' = (x^L - x)/(x^L - x^S)$, was used to normalize the R value, since it negatively impacted CLV. x' and x represented the normalized and original R (F, M) values, while x^L and x^S represented the largest and smallest R (F, M) value of all customers. After clustering customers's data to segment them based on RFM model, we also propose customer retention strategies after segmenting customer base (Fig. 1).
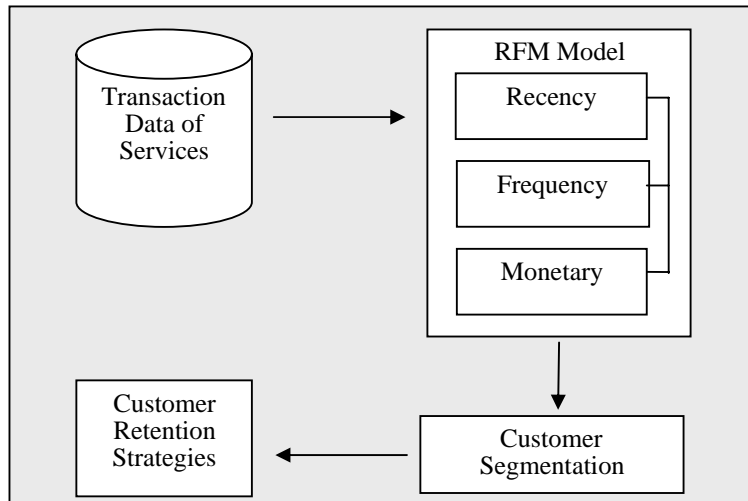
Fig. 1 Conceptual framework of research

4. Empirical Results

Table 2 presents the clustering results, listing eight clusters, each with the corresponding number of customers and their average R, F and M values. The last column also shows the overall average for all customers. These, for each cluster, were compared with the overall averages. If the average R (F, M) value of a cluster exceeded the overall average R (F, M), then H (High) was included and otherwise L (Low) was assigned. Two last row of the table show the RFM pattern for each cluster and the number of cases in each cluster respectively. Table 3 shows distances between final cluster centers, these results can be used only for descriptive purposes. The numbers show distance between a cluster and other ones that computed based on distance between cluster centers.

Table 2: Final Cluster Analysis

	Cluster								average
	1	2	3	4	5	6	7	8	
Recency	23	81.48	41	58.68	72.86	75	5.33	54.39	51.47
Frequency	12	22.6	26.67	10.14	57.49	12	10.33	78.61	28.73
Monetary	15000	905.96	173.66	1658.93	543.71	15000	1956	351.44	4448.71
Cluster Type	L, L, H	H, L, L	L, L, L	H, L, L	H, H, L	H, L, H	L, H, L	H, H, L	
Number of Cases	3	50	6	28	84	22	3	18	

Table 3: Distances between Final Cluster Centers

Cluster	1	2	3	4	5	6	7	8
1		1.158	1.024	.978	1.237	.591	.893	1.297
2	1.158		.465	.301	.418	.951	.880	.721
3	1.024	.465		.295	.510	1.075	.463	.623
4	.978	.301	.295		.578	.909	.607	.802
5	1.237	.418	.510	.578		1.100	.948	.323
6	.591	.951	1.075	.909	1.100		1.176	1.269
7	.893	.880	.463	.607	.948	1.176		.976
8	1.297	.721	.623	.802	.323	1.269	.976	

Then, discriminant analysis is used to determine whether clusters could be used to distinguish the sample customers (whether statistically significant). The analysis rejected the null hypothesis H_0 because the P-values were significant (P<0.05). The result confirmed that these clusters can significantly distinguish sample customers (Table 4). Table 5 also shows Eigenvalues, as first column depicts, three canonical discriminant functions were used in the analysis and Eigenvalues are significant.

Table 4: Wilks' Lambda test

Step	Number of Variables	Lambda	df1	df2	df3	Exact F				Approximate F			
						Statistic	df1	df2	Sig.	Statistic	df1	df2	Sig.
1	1	.034	1	7	206	830.060	7	206	.000				
2	2	.003	2	7	206	507.037	14	410	.000				
3	3	.000	3	7	206					400.53	21	586.33	.000

Table 5: Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	40.181	72.9	72.9	.988
2	10.619	19.3	92.2	.956
3	4.317	7.8	100.0	.901

Figure 2 shows group centroids graphically; at this figure the location of each cluster center has been depicted. As shown, X and Y axis are function 1 and 2 amounts respectively.
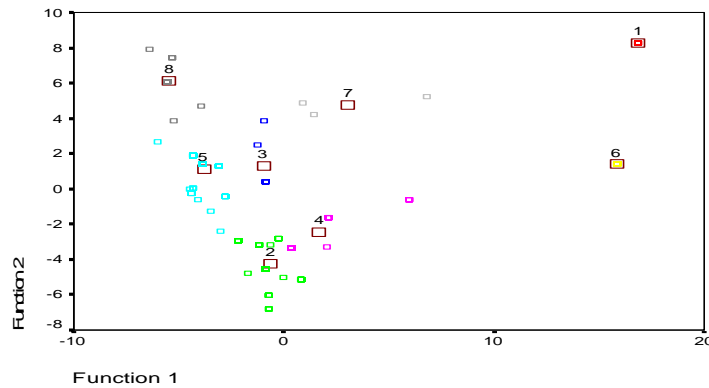
Fig. 2 Cluster number of cases

Finally, table 6 shows classification results; at each row, predicted group membership for each cluster has been showed, as seen, all cases but case 5 predicted to assign to only one cluster (100% fit), so according to calculations, 99.5% of original grouped cases correctly classified.

As mentioned earlier, each cluster represents a market segmentation. Customers in clusters with the pattern R (L) F (H) M (L) are considered to be loyal but not high valuable, purchased recently, purchase frequently, but spend not regularly with the bank. They are silver (not gold) customers. Bank should satisfy their needs perfect and provide various services to match with their needs to promote transaction values. Clusters with the pattern R (L) F (L) M (L) may include new customers who have only recently visited the bank. Customers in such clusters may be trying to develop closer relationships with the bank. These customers may become gold customers. Clusters with the pattern R (L) F (L) M (H) and also the pattern R (H) F (L) M (H) may include high valuable customers but not loyal to this bank, they should be treated specially to transact more. Finally, clusters with the pattern R (H) F (L) M (L) or the pattern R (H) F (H) M (L) include those who very rarely visited the site and made very few or low value transactions. They are valueless customers and bank should provide extra services to attract such customers.

Table 6: Classification Results

Cluster Number of Case	Predicted Group Membership								Total	%
	1	2	3	4	5	6	7	8		
1	3	0	0	0	0	0	0	0	3	100
2	0	50	0	0	0	0	0	0	50	100
3	0	0	6	0	0	0	0	0	6	100
4	0	0	0	28	0	0	0	0	28	100
5	0	1	0	0	83	0	0	0	84	98.8
6	0	0	0	0	0	22	0	0	22	100
7	0	0	0	0	0	0	3	0	3	100
8	0	0	0	0	0	0	0	18	18	100

5. Conclusions and further research

As marketing strives to become more accountable, we need metrics and models that help us assess the return on marketing investment. Many CRM researches pertain to develop a comprehensive model of customer profitability since the question 'Who are profitable customers?' is a starting point of CRM. CLV is one such metric. Many models have been researched to calculate CLV of a customer. The easy availability of transaction data and increasing sophistication in modeling has made CLV an increasingly important concept in both academia and practice. In this paper, we suggested a CLV model considering the recency, frequency and monetary at the same time. It clusters customers into segments according to their lifetime value expressed in terms of RFM. Moreover, clustering customers into different groups helps decision-makers identify market segments more clearly and thus develop more effective strategies. The authors recommend researchers to work on more measures and consider it comprehensively or rather compare various CLV models in a specific industry.

Acknowledgement

References

1. Amir, E. and Lev, B. (1996). "Value- relevance of nonfinancial information: The Wireless Communications Industry", Journal of Accounting and Economics, 22, pp.3- 30.

2. Balachandran, S. (2005). "How to build a culture of cost management",

Financial Times, p.16.

3. Bayo´n, T., Gutsche, J. and Bauer, H. (2002). "Customer equity marketing: touching the intangible", European Management Journal, 20(3), pp.213- 222.

4. Berger, P. D. and Nasr, N. I. (1998). "Customer lifetime value: marketing models and applications", Journal of Interactive Marketing, 12(1), pp.17- 30.

5. Bitran, G. R. and Mondschein, S. (1996). "Mailing decisions in the catalog sales industry", Management Science, 42(9), pp.1364- 1381.

6. Blattberg, R.C. and Deighton, J. (1996). "Manage marketing by the customer equity test", Harvard Business Review, pp.136- 144.

7. Brookes, R. and Palmer, R. A. (2004). The new global marketing reality, Basingstoke' Palgrave.

8. Buckinx, W. and Van den Poel, D. (2005). "Customer base analysis: Partial defection of behaviorally- loyal clients in a non- contractual fmcg retail setting", European Journal of Operational Research, 164 (1), pp.252- 268.

9. Chan, C. C. H. (2005). "Online auction customer segmentation using a neural network model", International Journal of Applied Science and Engineering, 3(2), pp. 101- 109.

10. Colombo, R. and Jiang, W. (1999). "A stochastic RFM model", Journal of Interactive Marketing, 13(3), pp. 2- 12.

11. Cooper R. and Kaplan R. S. (1988). "Measure costs right: make the right decisions", Harvard Business Review, 66, pp. 96- 103.

12. Courtheoux, R. (1995). Customer retention: how much to invest. Research and the Customer Lifecycle, New York, NY: DMA.

13. Denison, T. and McDonald, M. (1995). "The role of marketing: Past, present and future", Journal of Marketing Practice, 1(1), pp. 54- 76.

14. Dwyer, F. R. (1997). "Customer lifetime valuation to support marketing decision making", Journal of Interactive Marketing, 11(4), pp.6- 13.

15. Fader, P. S.; Hardie, B. G. S. and Lee, K. L. (2005). "Counting your customers the easy way: An alternative to the Pareto/ NBD model" Marketing Science, 24 (2), pp. 275- 284.

16. Goebel, D. J.; Marshall, G. W. and Locander, W. B. (1998). "Activity-based costing: accounting for a market orientation", Industrial Marketing Management, 27, pp. 397- 510.

17. Gro¨nroos, C. (2000). Service management and marketing, Chichester' John Wiley and Sons.

18. Gummesson, E. (1999). Total relationship marketing, Oxford' Butterworth- Heinemann.

19. Gupta, S. and Lehmann, D. R. (2003). "Customers as assets", Journal of Interactive Marketing, 17(1), pp.9- 24.

20. Gupta, S.; Hanssens, D.; Hardie, B.; Kahn, W.; Kumar, V.; Lin, N.; Ravishanker, N. and Sriram, S. (2006). "Modeling Customer Lifetime Value", Journal of Service Research, 9(2), pp.139- 155.

21. Hansotia, B. and Rukstales, B. (2002). "Incremental value modeling", Journal of Interactive Marketing, 16(3), pp.35–46.

22. Hoekstra, J. C. and Huizingh, E. K. R. E. (1999). "The lifetime value concept in customer- based marketing", Journal of Market Focused Management, 3(3- 4), pp. 257- 274.

23. Hsieh, N. C. (2004). "An integrated data mining and behavioral scoring model for analyzing bank customers", Expert Systems with Applications, 27, pp. 623- 633.

24. Hwang, H.; Jung, T. and 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, pp. 181- 188.

25. Jackson, D. R. (1994). "Strategic application of customer lifetime value in the direct marketing environment", Journal of Targeting Measurement and Analysis for Marketing, 3(1), pp. 9- 17.

26. 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- 45.

27. Jonker, J. J.; Piersma, N. and Poel, Dirk V. (2004). "Joint optimization of customer segmentation and marketing policy to maximize long- term profitability", Expert Systems with Applications, 27, pp. 59- 168.

28. Lindgreen, A. and Wynstra, F. (2005) "Value in business markets: What

do we know? Where are we going?" Industrial Marketing Management, 34(7), pp.732- 748.

29. Morgan, R. M. and Hunt S. D. (1994). "The commitment–trust theory of relationship marketing", Journal of Marketing, 58, pp. 20- 38.

30. Parvatiyar, A. and Sheth, J. N. (2000). "The domain and conceptual foundations of relationship marketing" In Sheth, J.N. and Parvatiyar A. (Eds.), Handbook of relationship marketing, Thousand Oaks, California' Sage Publications, pp. 3- 38.

31. Pearson, S. (1996). Building brands directly: creating business value from customer relationships, London: MacMillan Business.

32. Pfeifer, P. E. and Carraway, R. L. (2000). "Modeling customer relationships as Markov Chains", Journal of Interactive Marketing, 14(2), pp.43- 55.

33. Reichheld, F. F. and Sasser, E. W. (1990), "Zero defections: quality comes to services", Harvard Business Review, 68, pp. 105- 111.

34. Roberts, M. L. and Berger, P. D. (1989). Direct marketing management, Prentice- Hall: Englewood Cliffs, NJ.

35. Stahl, H. K.; Matzler, K. and Hinterhuber, H. H. (2003). "Linking customer lifetime value with shareholder value", Industrial Marketing Management, 32(4), pp.267- 279.

36. Tapscott, R. and Caston, A. (1993). Paradigm shift, New York' McGraw- Hill.

37. Verhoef, P. C. and Donkers, B. (2001). "Predicting customer potential value an application in the insurance industry", Decision Support Systems, 32, pp.189- 199.

38. Webster, J. F. E. (1992). "The changing role of marketing in the corporation", Journal of Marketing, 56(4), pp. 1- 17.

39. Yao, J.; Li, Y. and Chew, L. T. (2000). "Option price forecasting using neural networks", The International Journal of Management Science, 28, pp. 455- 466.