Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia
Social and Behavioral Sciences
ELSEVIER
Procedia - Social and Behavioral Sciences 62 (2012) 766 – 774

WC-BEM 2012

Investigating two customer lifetime value models from segmentation perspective

Abdulkadir Hiziroglu [a], Serkan Sengul [b]

[a] Assistant Professor at Sakarya University Industrial Engineering Department, Esentepe Campus, 54187, Sakarya, Turkey
[b] Undergraduate Student at Sakarya University Industrial Engineering Department, Esentepe Campus, 54187, Sakarya, Turkey

Abstract

Customer lifetime value has been a topic of interest for some years upon which plenty of academics and marketing managers have been dwelling. The topic plays a key role in customer relationship management and has been implemented in variety of sectors. The main goal of customer lifetime value is to specify the importance level of each customer for a company. Such questions as what sort of marketing strategies should be preferred for which customers, how much investments should be made for them and which marketing campaigns should be followed can all be determined by calculating lifetime value of customers. Many researchers have proposed different types of models for calculating customer lifetime value. Yet, the related literature lacks of comparative research on assessing the existing models, especially within the scope of segmentation. This paper aims at providing a classification for the current models in the literature based on their basic characteristics and making a comparison between two representative models from different classes using the same database. An evaluation from segmentation perspective was done and the results indicated that the model that represented the future-past customer behaviour model class was found to be superior than its peer using the same database and variables. © 2012 Published by Elsevier Ltd.

Keywords: Customer lifetime value, customer relationship management, customer segmentation, customer lifetime value models;

1. Introduction

Customer lifetime value (CLV) is quantitative measurement of a firm's net cash flows generated by its customers throughout their relationship with the company. This measurement has been of great importance and widely used by a variety of companies such as financial institutions, retail stores, telecommunication firms and etc. to find the differences between the customers and to tailor the most appropriate services for them. Correct calculation of CLV can facilitate a firm classify its customers based on their lifetime value rankings so that different marketing strategies can be developed for each group (Gupta and Lehmann, 2003).

Various models have been proposed by many researchers in order to calculate CLV. These models determine the CLV for each customer by using the available data in firms' customer databases. Some of these models calculate the lifetime values by only using the past data of customers, while some others take the future behaviour of them into consideration (Kumar, 2005). The literature was dominated by the studies in the latter category of the models. However, the current work lacks of comparative research on assessing those models, especially within the context of segmentation (Lemon and Tanya, 2006).

The aim of this paper is to provide a classification for the current models in the literature based on their basic characteristics and to make a comparison between two representative models from two different categories using the same database within the scope of segmentation. The rest of the paper is organized as the followings. The classification of the existing models and the empirical studies of the related literature are provided in Section 2. Section 3 presents the methodology followed in this study. Calculation of lifetime values for each model and the segmentation structures obtained by the comparative models are given in Section 4. Section 4 also includes the

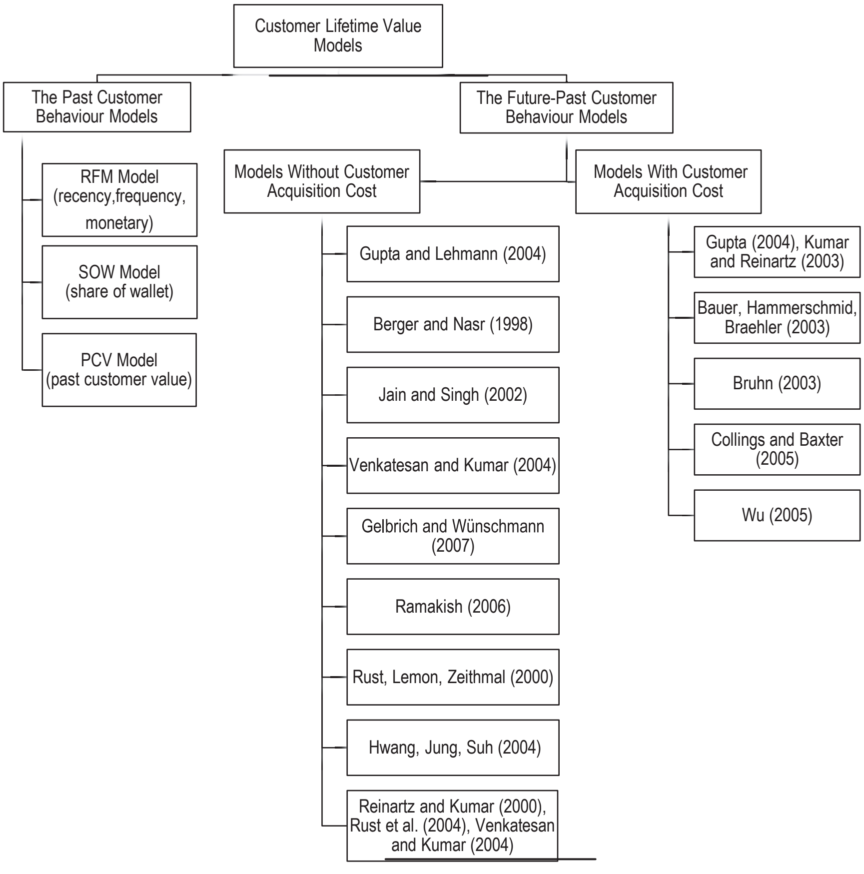
1877-0428 © 2012 Published by Elsevier Ltd. Selection and/or peer review under responsibility of Prof. Dr. Hüseyin Arasli
doi:10.1016/j.sbspro.2012.09.129

results of the comparison via testing the associated research hypotheses. Section 5 concludes the paper with information regarding practical and academic contribution of the study as well as its limitations.

2. Literature Review

Several models have been proposed in modelling CLV. In general, these models can be classified into two groups: past customer behaviour models and future-past customer behaviour models. There are mainly two differences between the future-past customer behaviour and the past customer behaviour models. The first difference is based on the assumption that if the customers subject to assessments will be active in the future. The second difference is that whether costs of customers are included into the models or not. The first group of models makes the calculations by including future activation rate of customers and as well as costs associated with the customers, while the latter group does not take these into account. Future-past customer behaviour models may also be separated into two categories based on the attribute of whether they include customer acquisition cost or not (See Table 1).

Table 1. Classification of CLV Models



Each model in the past customer behaviour group possesses certain parameters unique to the model's characteristics. RFM model is the most widely used method among them and it has been utilized in direct marketing area for around 30 years (Gupta et al., 2006). The models in the future-past customer behaviour category share the same underlying principle that for every customer, first how long it will be active is determined then net present values of these customers were calculated throughout the activation period. It is possible to find empirical studies in the related literature that utilized one of the above mentioned models as illustrated in Table 2.

When the existing empirical studies are reviewed, 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 segmentation. This is also highlighted a study in which projections on the future of CLV are provided (Lemon and Tanya, 2006). 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 body of the literature by providing the results of an empirical work conducted on two representative models in a comparative framework with a special focus on segmentation.

3. Methodology

In this study, based on the classification provided in the previous section two representative models from the groups of models will be compared within the context of segmentation. In order to accomplish that the variables in the acquired database 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 basic structural model were chosen for comparison as they both need the same set of variables.

3.1. Research Questions and Hypothesis

Customer segments will be generated for both of the models according to lifetime value assessment of each customer obtained by those representative models. Segmentation structure of each model will then be compared in order to determine which model is superior to its peer. Therefore, the answers of the following questions will be investigated:

1. To what extent the segments obtained by each model are appropriate?
2. Do both segmentation structures differ from each other?
3. Which segment structure should be chosen?

The above research questions can be expressed via the following research hypotheses:

$H1_0$: There are no statistically significant differences between the segments generated by RFM model.

$H2_0$: There are no statistically significant differences between the segments generated by basic structural model.

$H3_0$: The segment structures of each model are not different.

$H4_0$: The segmentation generated via RFM model is better than the segmentation generated by basic structural model.

Table 2. Summary of Empirical Studies Related to CLV

Study	Industry	Model	Data	Main Objectives	Main Results
Kumar et al. (2008)	Information Technologies	Reinartz and Kumar (2000), Rust et al. (2004), Venkatesan and Kumar (2004)	The database, which includes the information of SMEs who works with IBM, consists of 20.000 customers. <u>Parameters:</u> Average cost for a single marketing contact, index for time period, observation time frame, monthly discount rate	Which customers should be targeted? How the sources have to be divided for the customers?	The marketing investments which are done after the segmentation showing which customers' values will increase in the future would contribute to firm very much.
Reinartz and Kumar (2000)	Retail	Berger and Nasr (1998) Model	It belongs to a firm in retail industry in US consisting of 4965 customers. <u>Parameters:</u> Gross contribution in month, mailing cost, discount rate	How the CLV related to the profitability? Does the profitability increase when the CLV increases?	In spite of some exceptions the customers whose CLV are higher profits more.
Liu and Shih (2004)	Retail	RFM	Provided from a firm working in industrial and medical retail area. <u>Parameters:</u> Last transaction date, frequency of transaction, revenue	To determine which product is proper for the different customers.	It was determined that after the segmentation processes, which product is good for each customer groups could be found.
Glady, Baesens and Croux (2009)	Banking	Berger and Nasr (1998) Model	Provided from a finance corporation in Belgium. It shows the information between 2000-2005 consisting of 460.566 customers <u>Parameters</u> Annual cash flow, discount rate	To show the benefit of new version of Pareto/NBD analysis by adding some stuff to former one.	The new approach of Pareto has higher performance than the older one.
Hwang, Jung and Suh (2004)	Telecommunication	Hwang, Jung, Suh (2004) Model	Provided from a telecommunication firm in Korea. There are 200 different variables of 16.384 customers. For this study only 2000 of them was selected. <u>Parameters</u> Service period index, total service period, expected service period, past, future and potential contribution of customer.	To find a new CLV method and generate a segmentation depending of the results of the new model.	A new CLV model was proposed containing the customers' past and potential contribution and a segmentation was generated.
Gloy, Akridge and Preckel (1997)	Petroleum	Berger and Nasr (1998) Model	Provided from a firm which sells petroleum in an urban area consists of 3.281 customer's informations. <u>Parameters</u> Annual cash flow, discount rate	To show the importance of CLV in order to decide in petroleum marketing area.	CLV has a key role to determine marketing strategies, because it attends to show the requirement of considering the long term customer behaviour investigating.

3.2. Research Procedure

The methodology followed in this study is illustrated in Figure 1:

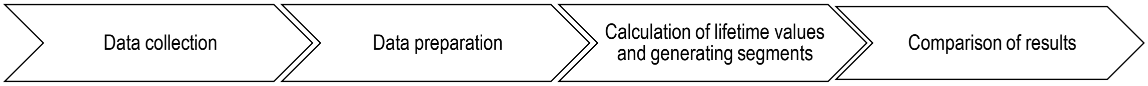


Figure 1. Study Methodology

Data Collection: The data was procured from www.dbmarketing.com website and the database belongs to a firm which engages in catalogue marketing in the USA. It consists of 1487 customer's information between the years of 2003-2007. Such variables as customer number, name, surname, address, city, state, zip code, last purchase date, number of orders and revenue are included in the database.

Data Preparation: Both of the representative models were performed by considering the last purchase date, the number of orders and revenue variables. The operationalisation of these variables for each model is provided in Table 3 and 4.

Table 3. Operationalisation of the variables for RFM model

Variable	Explanation	Operationalisation
R	Duration between a customer's last purchase date and present time	The present time was presumed as 01.01.2011. This date and the customers' last purchase dates were transformed into number versions in Excel and the numeric differences between them were accepted as recency value of those customers.
F	The number of transactions throughout the lifecycle of a customer	The total number of orders given by a customer was counted into a single value.
M	The revenue that is gained from a customer during lifecycle	The revenues of customers were assigned as their monetary values.

RFM model is based on the past customer purchase behaviour and R, F, M notations indicate Recency, Frequency and Monetary values, respectively. In basic structural model, CLV is calculated as the sum of the net cash flows from the customers considering the time value of money throughout the expected life of customers (Jain and Singh, 2002). This model can be expressed as in the following formula;

$$CLV=\sum_{i=1}^{n}\frac{Ri-Ci}{(1+d)^{i-0.5}} \qquad \textit{(Equation 1)}$$

Table 4. Operationalisation of the variables for Basic Structural Model

Variable	Explanation	Operationalisation
n	Expected life of a customer	$n=\frac{1}{1-r}$ (Reicheld, 1996) value depends on the retention rate of customer. The retention rate (r) was calculated via the following formula: $r=(\frac{T}{N})^{k}$ (Kumar, 2005); where, T: The time elapsed between the acquisition year and the last year of purchase for customers. N: The time elapsed between the acquisition year and the present year. k: The number of orders made in the observation period.
C_i	Total cost of customer in period i	$C_i= T_i+ W_i$; where, T_i *(Transportation Cost):* The cost delivering the order, which is based on the distance (presumed 0.01$/mile). The distribution center was assumed to locate in New York. The transportation cost for a customer was calculated via considering the approximate distance between the customer's residential state available in the database and the state of distribution center. W_i *(Weight Cost):* The cost that is paid for the weight of the order (presumed as 0.90$/kg). The weights of the orders presumed as uniformly distributed between 1-5 kg.
R_i	Total revenue of customer in period i	The revenues of customers were assigned as their monetary values.
d	Discount rate (annual)	Assumed to be %30.

4. Empirical Results

4.1. Lifetime Value Assessment and Segmentation

Lifetime value assessments or calculations of all customers were first carried out and then the corresponding segments based on these values were established. With regards to RFM model, all customers were labelled according to their R, F, and M values using the operationalisation given in Table 3. 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. Thus, eight different R-F-M combinations were obtained in order to develop customer segments. These combinations were then classified into four groups according to their R, F and M status. Four obtained segments and their descriptions together with number of customers in each and the corresponding R-F-M combinations are provided in Table 5.

Table 5. Customer Segments for RFM Model

Segment Number	Description of the Segment	Number of Customers	R-F-M Combination
1	Valuable Customers	185	RL, FH, MH
2	New Customers	285	RH, FH, MH
			RL, FL, MH
			RH, FL, MH
3	Vulnerable Customers	52	RL, FH, ML
			RH,FH, ML
4	Valueless Customers	965	RH, FL, ML
			RL, FL, ML

As far as basic structural model is concerned, lifetime value of each customer was calculated using *Equation 1* and the operationalisation provided in Table 4. The customers were then ranked in descending order based on these values. The total number of segments in basic structural model was set equal to the segment structure obtained by RFM model in order to attain an equivalent comparative base. Therefore, the first 185 customers in the ranking were described as “valuable customers”, the followed 285 of them as “new customers” and the next 52 of them as “vulnerable customers”. The remaining customers were assigned to “valueless customers” category.

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 both models. The following $H1_0$ and $H2_0$ hypotheses were tested in order to understand whether there are statistically significant differences between all corresponding segments generated by each comparative model or not.

$H1_0$: There are no statistically significant differences between the segments generated by RFM model.

$H2_0$: There are no statistically significant differences between the segments generated by basic structural model.

Testing the abovementioned hypotheses will ensure that the segments generated for each model can be identified according to the corresponding segmentation bases used during the segmentation process. Therefore, ANOVA tests were performed at 0.05 level of significancy for each segmentation structure and results were obtained as given in Table 6. Since the levels of significancy for all corresponding variables of each model are found to be less than 0.05 $H1_0$ and $H2_0$ were rejected. Thus, it can be said that the average values of these variables for the associated segments were statistically different from each other. In other words, the segments generated by RFM model are differentiable and this conclusion is also valid for the results of basic structural model.

Table 6. Descriptive Statistics and Result of ANOVA tests for both Models

Model	Variable	Segment1 (average)	Segment2 (average)	Segment3 (average)	Segment4 (average)	F	Sig
RFM Model	R	1242.30	1721.39	1436.42	1488.16	54.358	.000
	F	2.8	1.7	2	1	40.036	.000
	M	64.0919	49.6947	30.4231	19.9145	665.551	.000

Basic Structural Model	CLV	39.51	47.05	40.13	49.32	6.419	.000

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

The following hypothesis was tested in order to ensure that the segmentation structure of each model is different from the other.

$H3_0$: The segment structures of each model are not different.

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 structure obtained by two representative models and it was found to be 0.504. 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). A Kappa index value of 0.504 cannot be considered as a high level of agreement which yields rejection of $H3_0$. Therefore, it possible to distinguish or discern the segment structures of each model. Such differences would provide a base for further comparison of the models.

4.2.3. Comparison of the Models from Segmentation Perspective

The main goal of this study is to assess the comparative lifetime value models at segment level in order to find out which one is superior to the other via testing the following hypothesis.

$H4_0$: The segmentation generated via RFM model is better than the segmentation generated by basic structural model.

The comparison was performed based on average revenues of the segments. Table 7 provides that information for each individual customer segments of the comparative models.

Table 7. The Calculation of Average Revenues of Customer Segments for Both Models

			Average Revenue	
Segment Number	Number of Customers	Percent of Customers (%)	RFM Model	Basic Structural Model
1	185	12.4	64.56	78.89
2	285	19.2	50.19	39.64
3	52	3.5	30.87	32.29
4	965	64.9	19.62	19.91

As it can be seen from Table 7, the average revenues for Segment 3 and 4 of each corresponding model are rather approximate to each other. Therefore, it can be concluded that the segmentation structures obtained through both modes yielded similar results with regards to selection of vulnerable and valueless customer segments. Therefore, examining these segments would make no impact on noticing 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 basic structural model yields higher gain compared to the corresponding results of RFM model. On the contrary, when looking at the difference at Segment 2 basic structural model's average revenue seems lower in comparison with the associated results of its peer. Overall assessments of the differences indicate that the segmentation established by basic structural model could be seen more effective compared to RFM model. Because basic structural model seems to be more capable of enabling the assignment of the most valuable customers into the same segment. In other words, basic structural model has the ability to facilitate performing attraction of lucrative customers in one group and classifying the new customers in a lower value segment. For instance, assuming that a 10% return rate was expected for a specific marketing campaign, 78 dollars of revenue can be generated via basic structural model while 64 dollars can be summed through RFM model from 18 highest value customers of the corresponding segments (Segment 1). 12 dollars of difference indicates the superiority of the segmentation schema obtained by basic structural model relative to the schema created via RFM model. Therefore, the associated hypothesis ($H4_0$) regarding the comparison can be rejected based on this evaluation.

5. Conclusion

Ensuring an effective segmentation structure based on values of the customers and managing the relationships accordingly have been of crucial importance for many organisations. Although variety of CLV models have been proposed, the current literature lacks of studies concerning assessing the efficiency of different models in understanding the superiority of them relative to each other, particularly from the segmentation perspective. The aim of this paper is to assess different customer lifetime value models within the scope of segmentation. Within this context, first, different CLV models were reviewed and a classification according to their basic characteristics was provided. Then, two representative models (RFM and basic structural model) based on this classification were chosen and they were empirically evaluated on the same dataset. A comparison was made according to the segmentation structure obtained by each model. The results indicated that basic structural model was found to be superior to its peer using the same set of variables.

Implementation of CLV models in variety of sectors and a segmentation schema established according to CLV calculation could be of particular interest for marketing practitioners in developing different marketing policies and campaigns tailored to each customer segment. Within this context, this study contributes to the current body of the literature in two ways. First, a review was made on several proposed CLV models and a classification of these models was provided based on certain characteristics. Second, a comparison of two representative lifetime value models was carried out through taking into account the ability of these models to be used in establishing effective customer segments.

Despite the abovementioned contributions of the study there are few limitations worth mentioning which could make the external validity of the corresponding results questionable. First, this study was carried out on a specific database procured from a catalogue marketing company. Therefore, more work needs to be done through taking into account different data sets pertaining to other sectors or application domains. Second, the comparison was made on selected lifetime value models. Including different comparative studies using other lifetime value models should also be considered. Third, the current work was performed based on certain assumptions due to lack of specific customer-related data/information. A future work may be needed to relax these assumptions and to perform similar analyses via acquiring data sets that could be of more appropriate to real conditions. Fourth, average revenues of the segments were considered as an indicator in assessing the capability of the comparative models for establishing an effective segmentation structure. However, other measurements or indicators regarding building an effective segmentation can also be utilized in further studies.

References

Berger, P.D. & Nasr, N.I. (1998). Customer lifetime value: marketing models and applications. *Journal Of Interactive Marketing*, 12, 17-30.

Benoit, D.F. & Poel, D.V. (2009). Benefits of quantile regression for the analysis of customer lifetime value ın a contractual setting: an application ın financial services. *Expert Systems With Applications*, 36,10475-10484.

Chau, D.C.K., Xiu, L. & Ngai, E.W.K (2009). Application of data mining techniques ın customer relationship management: a literature review and classification. *Expert Systems With Application*, 36, 2592-2602.

Furness, P. (2001). Techniques for customer modeling in CRM. *Journal Of Financial Services Marketing*, 5, 293-307.

Faulkner, M. (2003). *Customer management excellence*. UK: John Willey & Sons, LTD (chapter 5).

Gupta, S. & Lehmann, R. (2008). Models of customer value. In B. Wierenga, (Ed), *Handbook Of Marketing Decision Models* (pp. 255-290) New York: Springer Science+Business Media, Llc.

Gupta, S. et al. (2006). Modeling customer lifetime value. *Journal Of Service Research*, 9, 139-150

Gharaee, R. (2009). Modeling CLV For Finacial Service. Master Thesis. Lulea University Of Technology, Porsön.

Glady, N., Baesens, B., Croux, C. (2007). A modified pareto/nbd approach for predicting customer lifetime value. *Expert Systems With Applications* ,36, 2062-2071

Hwang, H., Jung, T. & Suh, E. (2003). A LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry. *Expert Systems With Applications,* 26,181-188

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

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.

Kumar, V. & Reinartz, W.J. (2000). On the profitability of long life customers in a noncontractual settings: an empirical investigation an implications for marketing. *The Journal Of Marketing,* 64, 17-35

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., Venkatesan, R., Bohling, T. & Beckmann, D. (2008). The power of CLV: managing customer lifetime value at IBM. *Marketing Science,* 27, 585-599

Landis, J.R. &Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics,* 33, 159-174

Lemon, K.N. & Mark, T. (2006). Customer lifetime value as the basis of customer segmentation. *Journal Of Relationship Marketing,* 5, 55-69.

Lilien, G.L., Rangaswamy, A. & Bruyn, A.D. (2007). *Principles of marketing engineering*. Canada: Trafford Publishing (chapter 2)

Liu, D.R., Shih, Y.Y. (2004). Hybrid approaches to product recommendation based on customer lifetime value and purchase preference. *The Journal Of System And Software,* 77, 181-191

Liu, D.R. (2005). Integrating AHP and data mining for product recommendation based on customer lifetime value. *Information & Management,* 42, 387-400

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.

Rud, O.R. (2001). *Data mining cookbook, modeling data for marketing, risk, crm*. Canada: John Willey & Sons,Inc. (Chapter 12)