ICEB 2002 Proceedings

International Conference on Electronic Business (ICEB)

Winter 12-10-2002

Combining Clustering and MCDM Approach for Evaluating Customer Lifetime Value Ratings

Ya-Yueh Shih

Duen-Ren Liu

Gwo-Hshiung Tzeng

Combining Clustering and MCDM Approach for Evaluating Customer Lifetime Value Ratings

Ya-Yueh Shih [1] Duen-Ren Liu [2] Gwo-Hshiung Tzeng [3]
[1,2]Institute of Information Management, National Chiao Tung University
[1]Department of Information Management, Minghsin University of Science & Technology
[3]Institute of Management of Techology, National Chiao Tung University
Hsinchu, Taiwan
moonshih@iim.nctu.edu.tw dliu@iim.nctu.edu.tw

Abstract

Creating successful transaction actions to retain customers for future re-purchasing is extremely important in fiercely competitive environments. Moreover, different market strategies should be practiced for customers with different lifetime values and loyalty ratings. This work proposes a method, which combines clustering analysis and multiple criteria decision-making approach to evaluate customer lifetime value ratings, and construct the classification rules for individual clusters in market segmentation. An empirical case involving a hardware retailer is illustrated to show the usefulness for evaluating customer lifetime value ratings.

1. Introduction

A number of studies have discussed the evaluation of customer lifetime values (CLV) in terms of RFM (Recency, Frequency, and Monetary) [3][4][5][6][7][14][15][11]. Goodman [3] suggested that RFM method would avoid focusing on less profitable and instead allow these resources to be diverted to more profitable ones. From the behavioral perspective, the RFM measuring method is an important method for assessing the relationship between enterprise and customers.

Hughes [5] proposed a method for RFM scoring, which involves sorting the real data of RFM individual into 5 customer quintiles. Meanwhile, Stone [15] hypothesized that different weights should be assigned to RFM variables according to industry characteristics. To analyze the value of customers who paid using credit cards, Stone suggested placing the highest weighting on the number of purchases, followed by the period of purchase time, while placing the lowest weighting on the amount of purchases. Although various combinations and weightings have been proposed, judging and weighting RFM variables remains subjective.

This study uses an analytic hierarchy process (AHP) [12][13] to evaluate the importance (weight) of each RFM variable by the perception of decision makers. The K-means clustering is used to group customers into those with similar lifetime values or loyalty based on the performance value of RFM. Each target market can be further ranked using multiple criteria decision making (MCDM) approach. Finally, classification rules are determined for each cluster using the decision tree algorithm C5.0. The classification rules are employed to predict other potential customers regarding their groups and loyalty ranking. The case of hardware retailer is used for the sake of illustration.

The remainder of this study is oraganized as follows. Section 2 reviews related work in evaluating CLV. Section 3 then outlines the methodology used herein. Next, Section 4 presents a case study for evaluation. Section 5 construct rulesets for target market using a decision tree approach. Finally, the conclusion summarizes the contributions of this study and outlines areas for further research.

2. Related work

2.1. Market segmentation

Chen et al. [2] noted that clustering is one of the data mining tools used to discover knowledge processes. Clustering aims to maximize variance among groups while minimizing variance within groups. In clustering, many algorithms have been developed, such as k-means, hierarchical, fuzzy c-means approaches and so on.

This study applies K-means method to cluster customers based on their CLVs. The K-means method involves iterative improvement that can compensate for a poor initial partition of data. The number of clusters must be predetermined with the K-means method.

2.2. Evaluation of CLV

Bult et al. [1] explain the RFM terms as follows: (1) R (recency): time period since the last purchase, and the lower the value is, the higher the probability of the customer making a repeat purchase; (2) F (frequency): number of purchases made within certain time period; higher frequency indicates higher loyalty; (3) M (monetary): the amount of money spent during a certain time period; increasing monetary contribution from customers indicates increased focus on the company supplying the products.

Hughes [5] developed a widely used method for evaluating RFM. Each RFM is divided into 5 quintiles. With the ordering of customers from the top to the bottom, in the increasing order of recency values, the R scores of top 20% customers are set to 1, while those of the bottom 20% customers are set to 5. The F scores and M scores are assigned similarly, using the decreasing order of

The Second International Conference on Electronic Business
Taipei, Taiwan, December 10-13, 2002

frequency and monetary values, respectively. Accordingly, the RFM score of the best customer equals 111, while that of the worst equals 555. Different marketing strategies can thus be developed for different customers. Stone [15] hypothesized that the weights of RFM variables vary depending on industry characteristics. To analyze the value of customers with credit cards, Stone proposed that the frequency should be assigned the highest weighting, followed the recency, and finally, the monetary.

In practical applications, RFM variables need to measure different weights in different industries (e.g., [15]). However, Stone [15] determined the RFM weightings subjectively, without using a systematic approach or evaluation to determine the RFM weightings. This study employs AHP to evaluate each weight (relative importance) among RFM variables, and specifically asks decision makers to make intuitive judgments about ranking order to produce pairwise comparisons.

2.3. Decision tree

The earliest decision tree algorithm was extended from Concept Learning System (CLS) – Iterative Dichomizer 3 (ID3), which works by computing a metric known as the information gain ratio [8]. The spirit of decision tree is achieved by maximizing the information gain threshold at each node in the decision tree and the way to evaluate is based on classification validation. The ID3 has been refined into C4.5 by [9]. This study uses the latest C5.0 algorithm proposed in 1998, and the See5 software released by the company Rulequest [10].

3. Methodology for evaluating CLV ratings

This study proposes a method that combines clustering analysis and MCDM approach to evaluate the customer lifetime values or loyalty based on weighted RFM. Figure 1 shows the proposed methodology.

(1) Prepare a marketing database from enterprise; and remove nonsensical records such as those of customers who have purchase amount but never create any transactions. Next, the characteristics and purchase behavior of consumers are generated using simple statistics.

(2) Extract RFM variables for each customer to evaluate their lifetime values.

(3) Segment the market according to customer lifetime values by K-means; use the analysis of variance (ANOVA) to test whether RFM significantly discriminates against these market segments; and rank each market by MCDM with two models for comparison: one, without any preference, meaning equal weights among RFM, and the other with preferences, i.e., weight assessment is based on the AHP [12][13].

(4) Construct the classification rules for each market segmentation, and predict another potential customers by See5/C5.0.

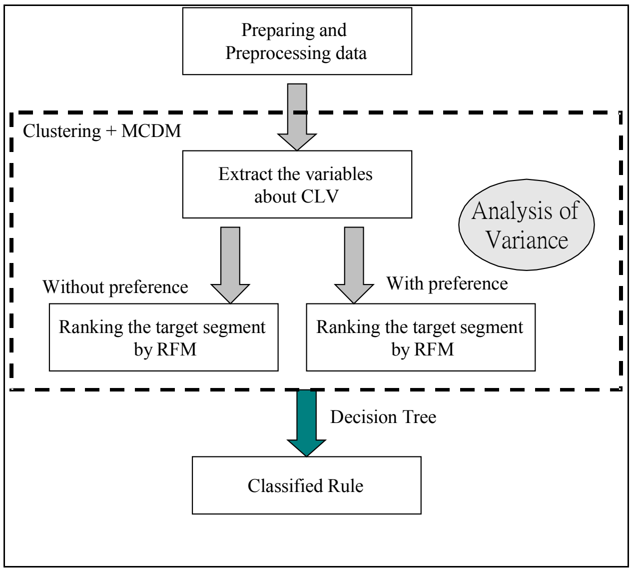

Fig. 1. Methodology for evaluating CLV ratings

4. Empirical study

An empirical case is used to show the usefulness of our approach for evaluating CLV ratings.

4.1 Preprocessing the data set

The empirical study is conducted using a data set collected from a company that manufactures wheels, casters, platform and hand trucks for industry, medical, hospital equipment and institutional purpose. The data set comprises 7,500 customers and 70,000 purchases records. These data were gathered from 2000/1 to 2002/4. Four related tables of transactions (Product, Transaction, Customer and Customer_classify) were extracted. After removing unreasonable records, the remaining contains 60,000 records.

Specifically, in the table Customer_classify, there are 3 types of customers, retailers, assembly industry and end users. Since the purchase records of retailers cover over 50% of all purchase records, this study analyzes the purchase records of retailers, which contains 984 customers.

4.2 Extracting the RFM values

Table 1 shows the RFM values for each 984 customers in hardware retailers, which are extracted from the dataset to measure CLV.

Table 1. RFM values for each customer

Customer no.	R	F	M
1108001	65	434	1252430
1108003	411	7	37930
…	…	…	…
1860003	159	87	313763

4.3 Evaluating CLV ratings with RFM clustering

4.3.1 Identifying target groups

The K-means method is used to group customers with similar lifetime value or loyalty. Eight possible combinations of inputs pattern (RFM) are made from 2 x 2 x 2. Table 2 shows the result. The average RFM variable values for each cluster are then compared with the total average RFM values of all clusters (163.57, 40.38 and 159481.1). If the average exceeds the total average, an upward arrow ↑ is given, while if the opposite occurs, a downward arrow ↓ is given.

Table 2. The result of K-means clustering

Cluster no.	Total Customers	Pattern
1	19	R ↑ F ↑ M ↑
2	136	R ↑ F ↓ M ↓
3	54	R ↑ F ↓ M ↓
4	30	R ↓ F ↑ M ↑
5	136	R ↓ F ↑ M ↑
6	467	R ↓ F ↓ M ↓
7	94	R ↑ F ↓ M ↓
8	48	R ↑ F ↓ M ↓

Analysis of variance is conducted to test whether RFM variables could discriminate 8 clusters. The analysis result would reject H_0, because the p-values are significant ($p < 0.05$). Thus, the result confirms that 8 clusters would be discriminated based on recency, frequency and monetary significantly.

Customers in cluster 4 and 5 have the same characteristics, both their average recency are less then the total average; frequency and monetary are greater than the total average (R ↓ F ↑ M ↑). Consequently, the customers in cluster 4 and 5 can be considered to be loyal customers who frequently visit and make large purchase.

Cluster 2, 3, 7 and 8 show the pattern of R ↑ F ↓ M ↓ , and are likely to be the clusters with customers of least loyalty. Such customers almost never visit and make transactions. Furthermore, they generally only make purchases during sales. Enterprises can reduce prices to attract these customers, but in so doing will suffer reduced margins. Cluster 6 displays the pattern of R ↓ F ↓ M ↓ and may represent new customers who have recently visited the company to make purchase.

Finally, Cluster 1 displays the pattern of R ↑ F ↑ M ↑ , and represents customers who may once have enjoyed a good relationship with the company, and have higher than average purchase frequency and purchase amount. However, customers in this cluster have not made transactions recently, possibly because of moving their businesses or bankruptcy.

4.3.2 Rating without preference on RFM

This section evaluates the rating on market segmentations, without preference on RFM variables, i.e., the RFM criteria have equal weight. Let $\boldsymbol{w} = [\boldsymbol{w_R}, \boldsymbol{w_F}, \boldsymbol{w_M}]$, where $\boldsymbol{w_R}$, $\boldsymbol{w_F}$, $\boldsymbol{w_M}$ represent the weight (relative importance) on RFM, respectively. The MCDM approach to evaluate the customer lifetime value ratings is illustrated as follows.

(1) *Determining the RFM weightings*
Herein, $\boldsymbol{w_R} = \boldsymbol{w_F} = \boldsymbol{w_M}$.

(2) *Constructing the normalized performance matrix*
The normalized performance matrix D contains 8 clusters associated with the RFM criteria. The 8 clusters are generated using K-means method based on RFM, as illustrated in Section 4.3.1. Let x_{ij} be the performance value of the ith cluster with respect to the jth RFM criteria in D. x_{ij} can be derived via computing the average R, F, and M value for each cluster j. Moreover, the normalized performance value (r_{ij}) of each cluster is derived as follows: the profit form, $r_{ij} = (x_{ij} - x_j^-)/(x_j^* - x_j^-)$, is used to normalize the frequency and monetary values, since they positively influence CLV or loyalty. The cost form, $r_{ij} = (x_j^* - x_{ij})/(x_j^* - x_j^-)$, is used for recency, since it has negative impact on CLV. Notably, x_j^* is the best performance value of all in jth criterion; x_j^- is the worst.

(3) *Ranking the performance order*
The weighted normalized performance value, $\boldsymbol{v} = [v_1, v_2, \ldots, v_8]^T$, can be derived by multiplying the normalized performance matrix $\boldsymbol{D}$ and the $\boldsymbol{w}^T$, i.e., $\boldsymbol{v} = [\boldsymbol{D} \times \boldsymbol{w}^T]$. The rankings of each market segmentation (cluster) can be determined according to the weighted normalized performance values of clusters, as shown in the right most column of Table 3. For example, the loyalty ranking of market segmentation 4 (cluster 4) equals 1.

Table 3. The rating of clusters without preference on RFM ($\boldsymbol{w_R}$, $\boldsymbol{w_F}$, and $\boldsymbol{w_M}$ are equal)

Cluster no.	R	F	M	Weighted Performance Value (v)	Loyalty Ranking
1	0.71	0.39	0.47	1.58	3
2	0.81	0.06	0.04	0.91	5
3	0.61	0.02	0.02	0.65	6
4	1.00	1.00	1.00	3.00	1
5	0.99	0.39	0.30	1.69	2
6	0.98	0.09	0.07	1.13	4
7	0.37	0.07	0.07	0.50	7
8	0.00	0.00	0.00	0.00	8
Average	0.68	0.25	0.25		

The fourth cluster displays the highest loyalty (Table 3), followed by the fifth cluster. That appears especially clear for decision makers to know which market segmentation is the main target of a company. Table 3 indicates that cluster 2, 3, 7 and 8 are the disloyal groups of customers, with group 8 being the worst, and thus the company should pay less attention to these groups and distribute less resource to them.

4.3.3 Rating with preference on RFM

This section conducts experiment to determine the ranking of CLV or loyalty for market segmentations (clusters) based on weighted RFM. AHP [12][13] is first used to assess the weightings (preferences) among RFM variables.

Three groups of evaluators exist: (a) 3 administrative department (managers); (b) 2 sales division (business manager, sales) and 1 marketing consultant; and (c) 5 customers who have made purchases. The above groups were invited to evaluate the criteria weightings. Data were gathered by interviewing evaluators, and an interview was conducted via a questionnaire (see Appendix, Table 5), with the answers being expressed in the form of a pairwise comparison matrix (see Appendix, Table 6).

(1) *Determining the RFM weightings*

According to the analytical result of AHP, $\boldsymbol{w_R}$, $\boldsymbol{w_F}$ and $\boldsymbol{w_M}$ are 0.7306, 0.1884 and 0.081, respectively. $\boldsymbol{w_R}$ has the highest ranking, followed by $\boldsymbol{w_F}$ and $\boldsymbol{w_M}$. The implication of this ranking on the preference of RFM is as follows. Recency is the most important, since the unit price of hardware products is relatively low, and thus evaluators only care about whether customers purchase continuously or not. In addition, customers, without any transaction activities in long periods, could have been lost or have transferred to new vendors.

(2) *Constructing the normalized performance matrix*

This step of constructing the normalized performance matrix $\boldsymbol{D}$ is the same as the step (2) in Section 4.3.2.

(3) *Ranking the performance order*

This step of ranking the performance order is similar to the step (3) in Section 4.3.2, except that $\boldsymbol{w_R}$, $\boldsymbol{w_F}$, and $\boldsymbol{w_M}$ are different, as derived from step (1). Table 4 shows the result of the rating on market segmentations, according to the preferences on RFM variables.

The ranking between cluster 1 and 6 in Table 4 differs from that in Table 3. That seems reasonable. For cluster 1 in Table 4, the frequency (0.39) and monetary (0.47) are higher than the total average F (0.25) and M (0.25). However, the time period since last purchase (recency) is very long, indicating that the customers in this cluster may have been lost or have transferred to other vendors.

For the cluster 6 in Table 4, although the frequency (0.09) and monetary (0.07) were lower than the total average, the recency (0.98) indicates that they have recently been active. Marketers should devote greater effort to retaining customers in cluster 6 than those in cluster 1. However, if individual RFM weightings are not considered, as the result shown in Table 3, cluster 1 ranks ahead of cluster 6. The comparison implies that the proposed approach may be a better method for evaluating the ranking of CLV, and the result is consistent with decision makers.

Table 4. The rating of clusters with preference on RFM ($\boldsymbol{w_R}$, $\boldsymbol{w_F}$, and $\boldsymbol{w_M}$ are different)

Cluster no.	R	F	M	Weighted Performance Value (v)	Loyalty Ranking
1	0.71	0.39	0.47	0.63	4
2	0.81	0.06	0.04	0.61	5
3	0.61	0.02	0.02	0.45	6
4	1.00	1.00	1.00	1.00	1
5	0.99	0.39	0.30	0.82	2
6	0.98	0.09	0.07	0.74	3
7	0.37	0.07	0.07	0.29	7
8	0.00	0.00	0.00	0.00	8
Average	0.68	0.25	0.25		

4.4 Classifying the target groups

See 5 software [10] based on C5.0 algorithm is used to construct decision tree and classification rules for each cluster. These classification rules are used to predict other potential customers belonging to which target market and mapping to which loyalty ranking. This approach can help decision makers to consider relevant marketing strategies. The samples are randomly divided into calibration (70%) and validation (30%) of total customers (984). The calibration samples are used as a training set to construct the classification rules, and then validation cases are classified for testing. The accuracy rate is critical to validating the classification result.

Each rule summarized the performance using the statistics (N/E, lift L) or (N, lift L) where: (1) N denotes the number of training cases covered by the rule; (2) E (if shown) represents the number of training cases covered that do not belong to the rule's class. Meanwhile, the accuracy of the rule is estimated by the Laplace ratio (N-E+1)/(N+2); (3) L is the estimated rule accuracy divided by the previous probability of the rule class. Taking rule 9 as an example, if Recency ≤ 117 and Frequency ≤ 63, then cluster 6 contains 325 customers. The accuracy of the rule is estimated to equal 0.997.

Rule 9: (325, lift 2.1)

Recency <= 117 Frequency <= 63 ➔ class 6 [0.997]

Finally, the classification result of the calibration and validation samples is evaluated. 11 rule sets are produced to classify 8 different clusters. The result shows 3.4% error rate of classification for validation samples.

5. Conclusions

In this paper, an analytical approach, combining clustering analysis and MCDM approach, is proposed to evaluate CLV ratings. The analytical result demonstrates that our approach would define the target market more clearly via AHP weighting and performance ranking than without weighting on RFM. The result helps market practitioners to make more effective strategies for retaining customers.

Moreover, to target potential customers, this study uses the classification approach to predict other potential customers for future purchases. Decision tree algorithm – C5.0 is used to classify these clusters (market segmentations) generated by K-means clustering. Classification rules based on RFM variables are extracted to classify 8 clusters. The result shows 3.4% error rate of classification for calibration samples.

There are two limitations of our study. First, we experimentally evaluate our approach on the data set collected from hardware retailers. Although customer purchase behavior are often available in marketing database, customer privacy and security concerns cause difficulty in obtaining more databases to verify whether our approach would be appropriate to other application domains, such as supermarkets or electronic commerce. Second, we assume that the relationship among RFM variables is linear. In fact, the purchase frequency always affects purchase amounts, and thus further research is required to relax the linear assumption.

Appendix

Table 5. AHP questionnaire sheet for RFM

	Importance degree									
Criteria	9	7	5	3	1	3	5	7	9	Criteria
		←						→		
Recency	9	7	5	3	1	3	5	7	9	Frequency
Recency	9	7	5	3	1	3	5	7	9	Monetary
Frequency	9	7	5	3	1	3	5	7	9	Monetary

Table 6. Retailer's RFM pairwise comparisons matrix

	Recency	Frequency	Monetary
Recency	1	5	7
Frequency	1/5	1	3
Monetary	1/7	1/3	1

References

[1] Bult, J.R., and Wansbeek, T.J. (1995), Optimal selection for direct mail, Marketing Science, 14 (4), pp. 378–394.

[2] Chen, M.S., Han, J., and Yu, P.S. (1996), Data mining: an overview from a database perspective, IEEE Transactions on Knowledge and Data Engineering, 8(6), pp. 866-883.

[3] Goodman, J. (1992), Leveraging the customer database to your competitive advantage, Direct Marketing, December, pp. 26-27.

[4] Ha, S.H., and Park, S.C. (1998), Application of data mining tools to hotel data mart on the Intranet for database marketing, Expert Systems with Applications, 15(1), pp.1-31.

[5] Hughes, Arthur M. (1994), Strategic Database Marketing, Chicago, Probus publishing.

[6] Kahan, R. (1998), Using database marketing techniques to enhance your one-to-one marketing initiatives, Journal of Consumer Marketing, 15(5), pp. 491-493.

[7] Miglautsch, J. (2000), Thoughts on RFM scoring, Journal of Database Marketing, 8(1), pp. 67-72.

[8] Quinlan, J.R. (1979), Discovering rules by induction from large collections of examples. In Expert Systems in the micro-electronic age, D. Michie (ed.), pp. 168-201. Edinburgh: Edinburgh University Press.

[9] Quinlan, J.R. (1993), C4.5: programs for machine learning, Morgan Kaufmann Publishers.

[10] Rulequest research, (1998), Data mining tools See5 and C5.0, http://www.rulequest.com/See5-infor.html.

[11] Saarenvirta, G. (1998), Data mining to improve profitability, CMA Magazine, March, pp. 8-12.

[12] Saaty, T.L. (1980), The analytic hierarchy process, McGraw-Hill, New York.

[13] Saaty, T.L., Vargas, L.G. (1982), The logic of priorities: application in business, energy, health, and transportation. Kluwer-Nijhoff, Boston.

[14] Shaw, M.J., Subramaniam, C., Gek Woo Tan and Welge, M.E. (2001), Knowledge Management and Data Mining for Marketing, Decision Support Systems, 31(1), pp. 127-137.

[15] Stone, Bob (1995), Successful direct marketing methods, Lincolnwood, IL: NTC Business Books.