Association for Information Systems
AIS Electronic Library (AISeL)

PACIS 2010 Proceedings

Pacific Asia Conference on Information Systems (PACIS)

2010

CRM Strategies for A Small-Sized Online Shopping Mall Based on Association Rules and Sequential Patterns

Beom-Soo Shim
Korea University, bumsuya@korea.ac.kr

Yong-Moo Suh
Korea University, ymsuh@korea.ac.kr

CRM STRATEGIES FOR A SMALL-SIZED ONLINE SHOPPING MALL BASED ON ASSOCIATION RULES AND SEQUENTIAL PATTERNS

Beom-Soo Shim, School of Business, Korea University, Seoul, South Korea, bumsuya@korea.ac.kr

Yong-Moo Suh, School of Business, Korea University, Seoul, South Korea, ymsuh@korea.ac.kr

Abstract

Data mining has a tremendous contribution to the extraction of knowledge and information which have been hidden in a large volume of data. This study has proposed customer relationship management (CRM) strategies for a small-sized online shopping mall based on association rules and sequential patterns obtained by analyzing the transaction data of the shop. We first defined the VIP customer in terms of recency, frequency and monetary value. Then, we developed a model which classifies customers into VIP or non-VIP, using various techniques such as decision tree, artificial neural network and bagging with each of these as a base classifier. Last, we identified association rules and sequential patterns from the transactions of VIPs, and then these rules and patterns were utilized to propose CRM strategies for the online shopping mall.

Keywords: Data mining, CRM strategy, Association rule, Sequential pattern, RFM.

1 INTRODUCTION

Internet technologies provide many competitive advantages such as agility, selectivity, individuality and interactivity (Weiber et al. 1998). The Internet enables customers to search products and services meeting their needs with less time than before. Online shopping malls, where buyer place orders over the internet, have emerged to become a prevalent sales channel. After Internet bubble in 2002 was broken, an uncountable number of small-sized online shopping malls have emerged every day due to many good characteristics of online marketplace, including significantly reduced search costs for products or services.

However, all of small-sized online shopping malls have not flourished. Many of them even vanished. As such, owners of the online shopping malls are concerned about how to manage the shops successfully. So is the owner of a small-sized online shopping mall in Korea, which deals with imported cute items such as those used in living room, kitchen, bathroom, etc. Until now since it started, the online shopping mall shows increased revenue for a while. But, these days, the revenue curve changes from steep to gentle slope. We suggested the owner of the online shopping mall that he needs good CRM strategies for the shop to thrive, as many other dot com companies have tried to develop CRM strategies that fit to them.

The objective of this research is to propose CRM strategies for the small-sized online shopping mall. To that end, from a given collection of transaction data of the shop, we 1) defined VIP in terms of recency, frequency and monetary value; 2) developed a classification model which classifies customers into VIP and non-VIP, using data mining techniques such as decision tree, artificial neural network and bagging; 3) identified association rules and sequential patterns hidden in the transaction data solely of VIP customers; 4) proposed CRM strategies based on the identified rules and patterns.

The rest of this paper is organized as follows. Section 2 presents literature review about data mining techniques for CRM. Section 3 explains experiments conducted in this study with respect to dataset and research framework, followed by step by step explanation in detail. Section 4 describes the experimental results by comparing the classification models and by attempting to interpret the identified association rules and sequential patterns that deserve application to the shop environment. Lastly, Section 5 ends the paper with a conclusion.

2 LITERATAURE REVIEW

This study proposes strategies which can be helpful on customer relationship management (CRM) in a small-sized online shopping mall. In this section, previous CRM-related researches using data mining techniques are reviewed.

2.1 Data mining for customer relationship management

Since the early 1980s, the concept of customer relationship management has gained its importance in marketing domain. Although it is difficult to make a totally approved definition of CRM, we can generally describe it as a comprehensive strategy including processes of acquiring, retaining and partnering with selective customers to create value for both the company and customers. Many previous CRM-related researches have used data mining techniques to analyze and understand customer behaviours and characteristics, and most of these have worked well (Bortiz et al. 1995; Fletcher et al. 1993; Langley et al. 1995; Lau et al. 2003; Salchenberger et al. 1992; Su et al. 2002; Tam et al. 1992; Zhang et al.1999). They have shown that data mining techniques can be used to elicit untapped useful knowledge from a large customer data. This section reviews previous researches which utilize classification, and association rules and/or sequential patterns analyses for various tasks in CRM domain.

Classification tasks have been carried out for various purposes in CRM domain. For example, Kim et al. (2006) adopted decision tree to classify the customers and strategy development based on customer lifetime value. Hwang et al. (2004) used logistic regression to segment customers based on their customer loyalty. Yu et al. (2005) identified interesting visitors through web log classification. Dennis et al. (2001) proposed customer knowledge management framework based on K-means. Kim and Street (2004) proposed a system for customer targeting based on ANN and genetic algorithm. Kim (2006) used logistic regression and ANN for feature selection to predict churn. Baesens (2004) identified the slope of the customer lifecycle based on Bayesian network classifier.

Other data mining techniques that are useful for the analysis of customer data are association rules and/or sequential patterns analyses. Researches based on association rules and sequential patterns have been conducted for various purposes in CRM domain. For example, LarivièRe and Poel (2005) used sequential patterns to predict future complaint. Chiang et al. (2003) extracted sequential patterns to predict network banking churn. Adomavicius and Tuzhilin (2001) examined association rules for one to one marketing. Aggarval and Yu (2002) and Kubat et al. (2003) identified association rules from market basket analysis. Changchien et al. (2004) used both ANN and association rules from market basket analysis technique to develop on-line personalized sales promotion.

Tables 1 and 2 provide a summary of previous researches which performed classification, and association rules and/or sequential patterns analyses for various purposes in CRM domain, respectively.

Specific Task	Data Mining Techniques	Reference
Customer segmentation and strategy development based on customer lifetime value	Decision tree	Kim et al. (2006)
Customer segmentation based on customer value	Logistic regression	Hwang et al. (2004)
Identification of interesting visitors through web log classification	Decision tree	Yu et al. (2005)
Suggestion of customer knowledge management framework	K-means	Dennis et al. (2001)
Suggestion of a system for customer targeting	ANN and genetic algorithm	Kim and Street (2004)
Feature selection for predicting churn	Logistic regression and ANN	Kim (2006)
Identification of the slope of the customer-lifecycle	Bayesian network classifier	Baesens et al. (2004)

Table 1. Previous classification and/or prediction researches in CRM domain

Specific Task	Data Mining Techniques	Reference
Prediction of future complaint	Sequential pattern	LarivièRe and Poel (2005)
Network banking churn analysis	Sequential pattern	Chiang et al. (2003)
One to one marketing	Association rule	Adomavicius and Tuzhilin (2001)
Market basket analysis	Association rule	Aggarval and Yu (2002), Kubat et al. (2003)
On-line personalized sales promotion	ANN and association rule	Changchien et al. (2004)

Table 2. Previous association rule and/or sequential pattern researches in CRM domain

2.2 RFM model definition

The RFM analytic model is proposed by Hughes (1994) which differentiates important customers based on the values of three variables, i.e., recency (R), frequency (F) and monetary value (M). They are defined as follows:

- R refers to the time interval between the last consuming behavior and current.
- F refers to the number of transactions over a certain period of time.
- M refers to the amount of money spent on products or services.

RFM are very effective values for customer segmentation (Newell 1997). According to the literature (Wu et al. 2005), researches showed that the bigger the values of R and F are, the more likely the corresponding customers are to make a new trade with companies. Moreover, the bigger M is, the more likely the corresponding customers are to buy products or services with companies again.

Data mining researches have been carried out based on RFM values in CRM domain. For example, Hosseini et al. (2010) adopted K-means algorithm to classify the customer loyalty based on RFM values. Cheng and Chen (2009) used K-means and rough set theory to segment customer value based on RFM values. Chen et al. (2009) identified purchasing patterns based on sequential patterns.

Table 3 indicates a summary of previous data mining researches based on RFM values in CRM domain.

Specific Task	Data Mining Techniques	Reference
Classification of customer loyalty	K-means	Hosseini et al. (2010)
Segmentation of customer value	K-means and rough set theory	Cheng and Chen (2009)
A purchasing pattern segmentation of customer	Sequential pattern	Chen et al. (2009)

Table 3. Previous data mining researches based on RFM values

3 EXPERIMENTS

This section explains the experiments conducted in this paper. First, we introduce a small-sized online shopping mall and briefly explain its transaction data. Then, we provide our research framework, with step by step explanation. Our experiments are conducted using Weka 3.6 and SAS Enterprise Miner 9.1.

3.1 Dataset

3.1.1 Target online shopping mall

A small-sized online shopping mall[1] in Korea provided its transaction data for analyses. It distributes and sells items imported from other countries by online. The online shopping mall deals with such cute items that are used in living room, kitchen, bathroom, etc. Its revenue is approximately 10,000 dollars per a month and has 2,017 registered members, so far. The size and revenue have been grown rapidly, since the online shopping mall opened on August 2008. However, the owner of the online shopping mall has little knowledge about their loyal customers and their purchasing patterns. Without actionable knowledge related to CRM, such as who loyal customers are and what their purchasing patterns are, it might be difficult to expect continuous sale increase.

3.1.2 Data description

The entire data covers the period from August, 2008 till October, 2009. The dataset consists of five tables such as demographic data, bulletin, comment, order management, and order product. In demographic table, there is information about member ID, age, member name, address, the first day of registration of members. Bulletin table has information about names of the writer, the number of click on bulletin, etc. Comment table has similar information to bulletin table. It consists of names of commenter, the number of reading comments, etc. In order management table, there is information about member ID, order ID, delivery address, zip code, telephone number, e-mail address, pay method (card or cash), pay amount, shipping fee, and mileage used. Order product table consists of order ID, category goods, product names, and product price. Unfortunately, original tables do not have sufficient fields proper to use. Having discussed with the domain experts, we can gather only 9 attributes from those tables.

[1] URL here (http://welesfamily.com/)

3.2 Research Framework

Figure 1 depicts the framework of our research. After integrating the original tables, we preprocessed the data, built classification models, and identified association rules and sequential patterns. After that, we derived CRM strategies by examining the identified association rules and sequential patterns from the transaction data of VIP customers. Details of each step are described in below.

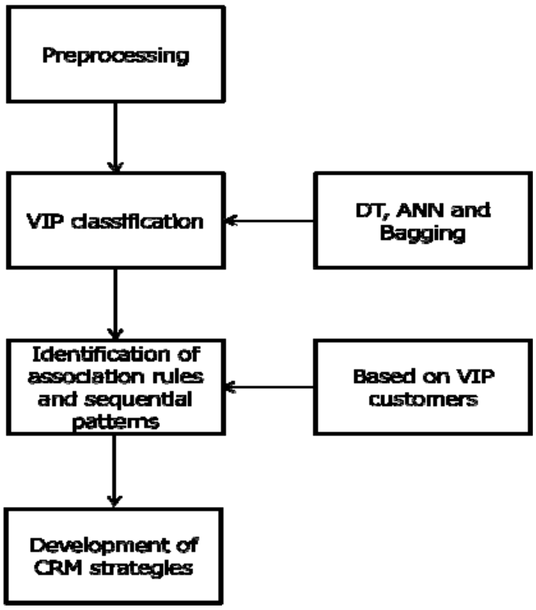

Figure 1. Framework for Experiments

3.2.1 Pre-processing

After integrating data tables, we deleted the duplicated records or those with many missing values and inaccurate values. Finally, we obtained 1,322 customers who have purchased products at the shop at least once, and 6,281 transaction data by them. Then, we defined VIP customers based on the values of RFM variables each customer has as follows:
We computed and normalized R, F and M values of 1,322 customers, sorted them in descending order, and selected the top X % of customers in each variable. We tried several times to compute the intersection of the three sets, with different values for X. When X is 75, the total monetary value of the customers (631 out of 1,322 customers) belonging to the intersection accounts for 70% of the total revenue of the shop. We defined the customers in the intersection as VIPs. Having discussed with the domain experts, we selected 9 more variables, besides R, F and M. Their meanings and data types are explained in Table 4. Examples of records of the resulting dataset are shown in Table 5.

Attribute	Meaning	Data type
R	The time interval between the last consuming behaviour and current	Real (0 to 100)
F	The number of transactions over a certain period of time	Real (0 to 100)
M	The amount of money spent on products or services	Real (0 to 100)
Age	Customer's age	Integer
reg_channel	Channel of registration (search or recommendation)	String
card_rate	Rate of paying in credit cards (ex. 0.25 means one use in four)	Real (0 to 1)
delivery	Delivery (a: 1 day delivery, b: 1 to 2 day delivery, c: 2 to 3 day delivery, d: more than 3 day delivery)	Character
comment	The number of posting comments on products	Integer
bulletin	The number of activities on bulletin board	Integer
reg_duration	The number of days after registration	Integer

mileage_used	The amount of used mileage in total	Integer
shipping_fee	The amount of paid shipping fee in total	Integer

Table 4. Independent attributes in this study

member_ id	age	reg_ channel	card_ rate	delivery	comment	bulletin	reg_ duration	mileage_ used	shipping_ fee
dlr4488	41	n	0	c	0	10	387	4,000	12,500
ejknock	36	search	1	a	0	1	62	0	0
hot1132	22	rec	0	c	0	1	52	0	0
jiuk486	33	n	1	c	3	5	248	13,410	0
luck777m	34	n	1	c	0	0	323	0	5,000

Table 5. Examples of records of our dataset

3.2.2 VIP classification

With the Weka 3.6 data mining tool which has been used widely, we developed models each of which classifies customers into VIP or non-VIP, using various data mining techniques such as decision tree, artificial neural network and bagging of each of these data mining techniques. DT and ANN are data mining approaches that have been heavily used for classification and/or prediction in support of marketing decision marking, and have shown good performance (Chien and Chen 2008; Kim et al. 2006; Yu et al. 2005; Kim and Street 2004; Kim 2006). As an alternative to a single classifier approach, bagging has been considered in the recent years (Frosyniotis et al. 2003; Kang and Doermann, 2003; Roli et al. 2004). The key idea of bagging is to combine a number of classifiers such that the resulting combined system achieves higher classification accuracy and efficiency than the original single classifiers. The aim to adopt bagging into this research is to compare with single classifier. C4.5 algorithm was used to build decision tree model with 0.25 confidence factor for pruning in this study. Having tried several times to conduct ANN, we set learning rate, momentum, epoch, and the number of hidden-layer to 0.1, 0.9, 50, and 1, respectively. In bagging method, we used 10 experts.

Prior to the model construction, we conducted feature selection by employing the wrapper approach with backward elimination. 12 input variables were evaluated using Gain Ratio attribute evaluator based on ranker search method, to select more influential variables when classifying the instances into VIP or non-VIP. The descending order of importance of 12 input variables is as follows: R, M, F, reg_duration, shipping_fee, age, comment, bulletin, card_rate, delivery, reg_channel, mileage_used. Comparisons of the classification results of models from 10-fold cross-validation are made in Section 4.

3.2.3 Identification of association rules and sequential patterns

In this study, we used the transaction data of only VIP customers. It contains 3,879 transaction data of 631 VIP customers. As shown in Table 6, each instance of transaction data consists of 5 attributes such as member_id, order_id, category, subcategory, and order_date. As Table 6, when customers purchase several items at the same time, instances exist as many as they purchase. We attempted to find association rules and sequential patterns among the categories as well as among the subcategories. Table 7 shows the category and subcategory of goods. Tables 9 and 10 show the results of association rules, and Tables 11 and 12 show the results of sequential patterns. After several attempts with different parameter values, we finally set the minimum support to 3.5% and the minimum confidence to 40% for categories analyses (Tables 9 and 11) and set the minimum support to 3% and the minimum confidence to 20% for subcategories analyses (Tables 10 and 12). To derive sequential patterns, we set the time window to the whole period of the dataset (15 months) because the number of instances was small and the whole period of our data was not so long.

Member_id	Order_id	Category	Subcategory	Order_date
angel9270	20090320-0000063	Dishes	bottle	2009-03-20

angel9270	20090331-0000235	Dishes	cup	2009-03-31
angel9270	20090331-0000235	Dishes	bowl	2009-03-31
angel9270	20090331-0000235	Living room goods	basket	2009-03-31
zzzmia	20090822-0000018	Dishes	tray	2009-08-22
zzzmia	20090822-0000018	Cooking goods	cooker	2009-08-22
zzzmia	20091101-0000038	Dishes	vessel	2009-11-01
zzzmia	20091101-0000038	Dishes	vessel	2009-11-01

Table 6. Examples of instances for association rules and sequential patterns

Subrogate of nomination	Category	Subcategory
A	Dishes	bottle, bowl, coaster, cup, dosirak, knife, plate, spoon, tray, vessel
B	Cooking goods	cooker, cutting board, kettle, pan
C	Kitchen goods	apron, napkin, oil paper, kitchen towel
D	Bathroom goods	body washer, toilet cover, basin, bath gloves, soap, tooth brush, bath towel
E	Living room goods	basket, air freshener, bag, blanket, box, calendar, case, clock, doll, handkerchief, hanger, mat, toy, pouch, purse, shoes
F	Washing goods / Cleansing supplies	broom, scouring pad, trash can
G	Office supplies	calculator, note, file, mouse pad, paste, pencil case, scotch tape

Table 7. Category and subcategory of goods

4 DISCUSSIONS

4.1 Results from classification models

12 experiments for each of the 4 data mining techniques, or a total 48 experiments in each dataset were conducted for classification analyses. In Figure 2 showing the classification accuracy as the number of input variables decrease, we can find out that the highest classification accuracy was acquired when we included the most influential 3 attributes such as R, F and M in our dataset as was expected. As shown in Table 8, DT (Decision Tree) and bagging-DT show the best performance among all data mining techniques. Although bagging-DT outperformed DT, the result of McNemar test applied to DT and bagging-DT with the most influential 3 attributes shows that the difference of accuracy between DT and bagging-DT is not statistically significant. Therefore, it is not necessary to use bagging-DT in order to classify customers into VIP or non-VIP, considering the time and efforts to build bagging models.

The number of features	DT	ANN	Bagging-DT	Bagging-ANN
12	97.12%	87.36%	97.65%	89.33%
11	97.12%	88.04%	97.65%	89.41%
10	96.82%	86.91%	96.89%	88.88%
9	96.82%	89.10%	97.12%	90.09%
8	97.12%	89.25%	97.12%	90.24%
7	96.97%	89.62%	97.20%	90.62%
6	96.97%	88.88%	97.35%	90.77%

5	96.97%	89.71%	97.50%	90.77%
4	97.12%	89.63%	97.42%	90.63%
3	97.12%	89.78%	97.65%	90.92%
2	92.13%	88.95%	92.13%	89.41%
1	72.61%	72.01%	73.07%	72.23%

Table 8. Accuracy comparisons of each mining technique

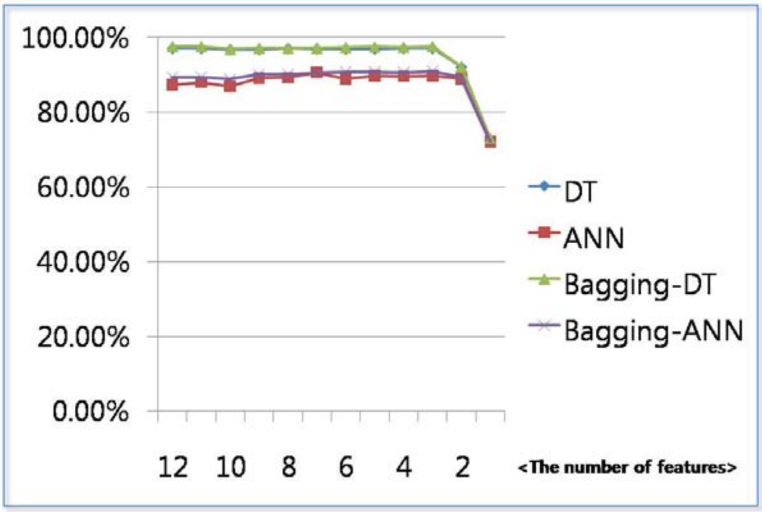

Figure 2. Classification accuracy of each mining technique

4.2 Results of association rules and sequential patterns

As shown in Tables 9 and 10, 17 association rules were found from VIPs. Also, in Tables 11 and 12, 14 sequential patterns were found from VIPs.

Table 9 shows that ‘E (Living room goods)’ is frequently associated with other categories. For example, people who purchase ‘A (Dishes)’, ‘C (Kitchen goods)’, or ‘D (Bathroom goods)’ generally tend to purchase each of them together with ‘E (Living room goods)’. Table 10 indicates association rules among subcategories. From Table 11, we can find out that customers who purchase category ‘A’ or ‘E’ are likely to buy goods of the same category later again. As shown in Table 12, customers who firstly purchase cup are likely to buy cup again, but those who first purchase bowl are likely to buy either cup or bowl later. It indicates that both cup and bowl are subcategories which have significant influence on revenue of the shop.

Lift	Support (%)	Confidence (%)	Rule
1.12	7.63	56.16	D => E
1.19	5.21	59.57	C => E
1.11	5.02	55.67	D & A => E
1.32	3.81	58.57	C & A => E
1.17	3.81	43.62	C => E & A

Table 9. The results of association rules on category

Lift	Support (%)	Confidence (%)	Rule
1.43	6.98	41.67	bowl => cup
1.43	6.98	23.89	cup => bowl
1.03	5.49	29.95	plate => cup

1.17	5.21	34.15	tray => cup
1.61	4.93	29.44	bowl => plate
1.61	4.93	26.90	plate => bowl
1.34	3.91	39.25	spoon => cup
1.05	3.72	30.77	case => cup
1.67	3.63	29.32	dosirak => bag
1.67	3.63	20.63	bag => dosirak
1.16	3.44	20.33	bottle => bag
1.44	3.07	25.38	case => bag

Table 10. The results of association rules on subcategory

Support (%)	Confidence (%)	Rule
11.57	41.48	A => A => A
8.24	43.33	E => A => A
7.29	63.01	A => A => A => A
7.29	40.00	A => E => A
6.97	47.83	E => E => A
5.86	48.05	A & E => A => A
5.86	40.22	E => E => E
5.55	67.31	E => A => A => A
5.39	73.91	A => E => A => A
5.07	45.07	A => A & E => A

Table 11. The results of sequential patterns on category

Support (%)	Confidence (%)	Rule
8.24	22.03	cup => cup
6.18	28.47	bowl => cup
4.91	22.63	bowl => bowl
3.33	40.38	cup => cup => cup

Table 12. The results of sequential patterns on subcategory

4.3 Development of CRM strategies

According to Swift (2001), Parvatiyar and Sheth (2001), and Kracklauer et al. (2004), CRM consists of four dimensions such as customer identification, customer attraction, customer retention, and customer development. Customer identification is meant to identify segments of potential customers, each of which includes customers who are relatively similar (Woo et al. 2005). Customer attraction attempts to attract the target customer segments by motivating customers to place orders through various channels (Cheung et al. 2003; He et al. 2004; Liao and Chen 2004). Customer retention refers to the activity of preventing the existing customers from switching to competitors by enhancing the level of customer satisfaction through one-to-one marketing, loyalty program, complaints management, etc. (Chen et al. 2005; Jiang and Tuzhilin 2006). Customer development, the ultimate goal of CRM, aims to maximize the revenue by expanding transaction intensity, transaction value and individual customer profitability through customer lifetime value analysis, up/cross selling and market basket analysis (Etzion et al. 2005; Rosset et al. 2003).

Based on these dimensions of CRM and the results obtained from this study, we decided to suggest target marketing strategies against the VIP customers for the online shopping mall as follows:

- The classification model can be made use of to identify VIP customers, so that the online shopping mall can exercise marketing activities against them. It will be more cost-effective than doing the same thing against all the customers. Of course, it will be necessary to develop a new classification model based on a new definition of VIP customer as more customer transaction data is accumulated. This cycle of collecting data, building models and using them for marketing activities should be continuous.

- Since category 'E' is highly associated with categories 'A', 'C', and 'D' as shown in Table 9, it is recommended that the online shopping mall redesign their web site so that the page associated with category 'E' can be one-click away from categories 'A', 'C' or 'D'.
- Although (cup => bowl) and (bowl => cup) are prevalent association rules and (cup => cup), (bowl => cup), and (bowl => bowl) are prevalent sequential patterns, the current keywords of the online shopping mall in major search sites do not include or imply cup and/or bowl. To attract prospective customers much more, it is highly suggested that the online shopping mall properly set keywords to those which imply cup and/or bowl with specific characteristics that are unique to the cups and bowls dealt with by the mall.
- Through the purchasing patterns in the form of association rules and sequential patterns found in this study, we can develop personalized promotion strategies. From Table 10, we suggest the following recommendation or bundling strategy:
  - We can recommend cups to the customers who purchased bowls. The same strategy can be applied to the other association rules among subcategories.
  - Especially, since (cup and bowl), (bowl and plate), and (dosirak and bag) have been sold at the same time, it is highly recommended that they are sold as a bundle of items.

  From Tables 11 and 12, we suggest the following strategies.
  - Having recognized the time gap between the first purchase and the second in a sequential pattern, we can send mails at the right time to the customers who purchased the item in the left hand side of the sequential pattern in order to remind them to buy the item in the right hand side of the sequential pattern.

5 CONCLUSIONS

In this study, we defined the VIP based on RFM values. Then, we developed classification models which classify customers into VIP or non-VIP using 4 data mining techniques and compared their hit ratios. Decision Tree and the bagging with decision tree show the highest hit ratios with neglectable difference. Then, we extracted 17 association rules and 14 sequential patterns from the transaction data of VIP customers, and suggested strategies for the small-sized online shopping mall based on these identified rules and patterns. It seems that the strategies suggested by this study can be utilized to enhance the revenue of the online shopping mall more efficiently than before.

This study has a few things to be desired. First, the target online shopping mall is of small size and of short history and thus the dataset we have used for the analysis is not big enough to derive more meaningful results. Second, for the same reason, the VIP customers we have selected account for almost half of the whole customers. The Pareto rule does not hold in this mall. Last, this study should have continued to see if the strategies derived from the actionable knowledge obtained by data mining techniques are effective or not. We plan to do it in the near future. Nonetheless, we believe that such experiments as conducted in this study deserve to be paid attention of small-sized online shopping malls where a huge amount of useful data still remains unused.

References

Adomavicius, G., and Tuzhilin, A. (2001). Expert-driven validation of rule-based user models in personalization applications. Data Mining and Knowledge Discovery, 5, 33–58.

Aggarval, C. C., and Yu, P. S. (2002). Finding localized associations in market basket data. IEEE Transactions on Knowledge and Data Engineering, 14, 51–62.

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, 508–523.

Bortiz, J. E., and Kennedy, D. B. (1995). Effectiveness of neural network types for prediction of business failure. Expert Systems with Applications, 9, 503–512.

Changchien, S. W., Lee, C. F., and Hsu, Y. J. (2004). On-line personalized sales promotion in electronic commerce. Expert Systems with Applications, 27, 35–52.
Chen, M. C., Chiu, A. L., and Chang, H. H. (2005). Mining changes in customer behaviour in retail marketing. Expert Systems with Applications, 28, 773–781.
Cheng, C. H., and Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. Expert Systems with Applications, 36, 4176–4184.
Cheng, Y. L., Kuo, M. H., and Tang, K. (2009). Discovering recency, frequency, and monetary (RFM) sequential patterns from customers' purchasing data. Expert Systems with Applications, 8, 241–251.
Cheung, K. W., Kwok, J. T., Law, M. H., and Tsui, K. C. (2003). Mining customer product ratings for personalized marketing. Decision Support Systems, 35, 231–243.
Chiang, D. A., Wang, Y. F., Lee, S. L., and Lin, C. J. (2003). Goal-oriented sequential pattern for network banking churn analysis. Expert Systems with Applications, 25, 293–302.
Chien, C. F., and Chen, L. F. (2008). Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry. Expert Systems with Applications, 34, 280–290.
Dennis, C., Marsland, D., & Cockett, T. (2001). Data mining for shopping centres customer knowledge management framework. Journal of Knowledge Management, 5, 368–374.
Etzion, O., Fisher, A., and Wasserkrug, S. (2005). E-CLV: A modeling approach for customer lifetime evaluation in e-commerce domains, with an application and case study for online auction. Information Systems Frontiers, 7, 421–434.
Fletcher, D., and Goss, E. (1993). Forecasting with neural networks: An application using bankruptcy data. Information and Management, 3.
Frosyniotis, D., Stafylopatis, A., and Likas, A. (2003). A divide-andconquer method for multi-net classifiers. Journal of Pattern Analysis and Applications, 6(1), 32–40.
He, Z., Xu, X., Huang, J. Z., and Deng, S. (2004). Mining class outliers: Concepts, algorithms and applications in CRM. Expert Systems with Applications, 27, 681–697.
Hosseini, M., Anahita, M., Mohammad, R., G. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. Expert Systems with Applications, 37, 5259–5264.
Hughes, A. M. (1994). Strategic database marketing. Chicago.
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, 181–188.
Jiang, T., and Tuzhilin, A. (2006). Segmenting customers from population to individuals: Does 1-to-1 keep your customers forever. IEEE Transactions on Knowledge and Data Engineering, 18, 1297–1311.
Kang, H.-J., and Doermann, D. (2003). Evaluation of the informationtheoretic construction of multiple classifier systems. In: Proceedings of the international conference on document analysis and recognition, Edinburgh, Scotland, Aug. 3–6, pp. 789–793.
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, 101–107.
Kim, Y. S. (2006). Toward a successful CRM: Variable selection, sampling, and ensemble. Decision Support Systems, 41, 542–553.
Kim, Y. S., and Street, W. N. (2004). An intelligent system for customer targeting: A data mining approach. Decision Support Systems, 37, 215–228.
Kracklauer, A. H., Mills, D. Q., and Seifert, D. (2004). Customer management as the origin of collaborative customer relationship management. Collaborative Customer Relationship Management - taking CRM to the next level, 3–6.
Kubat, M., Hafez, A., Raghavan, V. V., Lekkala, J. R., and Chen, W. K. (2003). Item set trees for targeted association querying. IEEE Transaction on Knowledge and Data Engineering, 15, 1522–1534.
Langley, P., and Simon, H. A. (1995). Applications of machine learning and rule induction. Communication of the ACM, 38, 55–64.

LarivièRe, B., and Poel, D. V. D. (2005). Investigating the post-complaint period by means of survival analysis. Expert Systems with Applications, 29, 667–677.
Lau, H. C. W., Wong, C. W. Y., Hui, I. K., and Pun, K. F. (2003). Design and implementation of an integrated knowledge system. Knowledge-Based Systems, 16, 69–76.
Liao, S. H., and Chen, Y. J. (2004). Mining customer knowledge for electronic catalog marketing. Expert Systems with Applications, 27, 521–532.
Newell, F. (1997). The new rules of marketing: How to use one-to-one relationship marketing to be the leader in your industry. McGraw-Hills, New York.
Parvatiyar, A., and Sheth, J. N. (2001). Customer relationship management: Emerging practice, process, and discipline. Journal of Economic & Social Research, 3, 1–34.
Roli, F., Kittler, J., and Windeatt, T. (2004). Multiple classifier systems. In: Proceedings of the 5th international workshop on multiple classifier systems, Cagliari, Italy.
Rosset, S., Neumann, E., Eick, U., and Vatnik, N. (2003). Customer lifetime value models for decision support. Data Mining and Knowledge Discovery, 7, 321–339.
Salchenberger, L. M., Cinar, E. M., and Lash, N. A. (1992). Neural networks: A new tool for predicting thrift failures. Decision Sciences, 23, 899–916.
Su, C. T., Hsu, H. H., and Tsai, C. H. (2002). Knowledge mining from trained neural networks. Journal of Computer Information Systems, 42, 61–70.
Swift, R. S. (2001). Accelarating customer relationships: Using CRM and relationship technologies. Upper saddle river. N.J.: Prentice Hall PTR.
Tam, K. Y., and Kiang, M. Y. (1992). Managerial applications of neural networks: The case of bank failure predictions. Management Science, 38, 926–947.
Weiber, R., and Kollmann, T. (1998). Competitive advantage in virtual markets: perspectives of information based Marketing. European Journal of Marketing, 32(7/8), 603–615.
Wu, C. H., Kao, S. C., Su, Y. Y., and Wu, C. C. (2005). Targeting customers via discovery knowledge for the insurance industry. Expert Systems with Applications, 29, 291–299.
Wu, J., and Lin, Z. (2005). Research on customer segmentation model by clustering. ACM International Conference Proceeding Series, 113.
Yu, J. X., Ou, Y., Zhang, C., and Zhang, S. (2005). Identifying interesting visitors through web log classification. IEEE Intelligent Systems, 20, 55–59.
Zhang, G., Hu, M. Y., Patuwo, B. E., and Indro, D. C. (1999). Artificial neural networks in bankruptcy prediction: General framework and cross validation analysis. European Journal of Operational Research, 116, 16–32.