

Impact Analysis of Data Mining on CRM: Model Integration and Analysis

Fatima Shoaib¹, Khuram Shafi¹, Zhou Dan², Supanida Muangna¹, Amna Nazeer³

¹School of Management, Huazhong University of Science and Technology, Wuhan, China

²Tongren Secondary Vocational School, Tongren Polytechnic College, Guizhou, China

³School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, China

Email address

shafikhuram@yahoo.com (K. Shafi), Iqtest00@hotmail.com (F. Shoaib)

To cite this article

Fatima Shoaib, Khuram Shafi, Zhou Dan, Supanida Muangna, Amna Nazeer. Impact Analysis of Data Mining on CRM: Model Integration and Analysis. *American Journal of Business, Economics and Management*. Vol. 3, No. 2, 2015, pp. 98-103.

Abstract

In this paper four dimensions of Customer Relationship Management (CRM) were identified. Of particular interest, however, it is potential link between Knowledge Management (KM), Direct Marketing (DM), Innovation and Technology. We examined the impact of four dimensions of CRM: Direct Marketing (DM) by organization, continuous improvement and innovation and knowledge management in organization, organization's technology adaption to change, and organization's management of environmental impacts. The relationship of these four of each dimension of CRM is investigated with a help of model. The multidimensionality of CRM was constructed is well acknowledged in literature. All dimensions of CRM and their relationship with four variables as we have mentioned above was briefly explained. In this research paper, positive relationship between internal and external Dimensions of KM, DM, IM, and TM is explained briefly.

Keywords

Management Information System, Data Mining, Customer Relationship Management, Business Innovation

1. Introduction

Customer Relationship Management (CRM) is a term borrowed from the information sciences for procedures, software and usually internet capabilities that help an organization to manage customer relationships in an organized way. CRM consists of a historical view and analysis that helps in reducing search and correlating customers and to foresee customer needs effectively and to increase business. The CRM integrated system is comprised of centralized system that may be used for attracting new customers and retaining the old one. Information sharing is the most important element for the real time implementation of CRM system, failing to which may lead towards the failure of the CRM system. CRM systems have many shortcomings and they need to overcome by using state of the art data mining techniques. Data mining is related to extract information from a data set and transform into understandable structure for further use.

Several researches have been conducted in analyzing the factors affecting the CRM systems. The number of customer relationship management (CRM) implementations has grown in recent years. However, few academic studies of the issues

associated with the implementation of the concept are available. CRM typically involving tracking individual customer behavior over time, and using this knowledge to configure solutions precisely tailored to the customers' and vendors' needs. In the context of choice, this implies designing longitudinal models of choice over the breadth of the firm's products and using them prescriptively to increase the revenues from customers over their lifecycle [1] [2]. It offers a modest contribution through the analysis of a case study of a CRM implementation at a UK-based manufacturing company. The case study illustrates that CRM is a complex and holistic concept, organized around business processes and the integration of information technologies. The study highlights that executing CRM requires effective leadership, sourcing, targeting and evaluation strategies [3]. In the effects of various dimensions of CRM on innovation capabilities are investigated, in which impact of CRM attributes (information sharing, customer involvement, long-term partnership, joint problem-solving and technology-based CRM) are tested on the five aspects of innovation capability (product, process,

administration, marketing and service innovations). While in [4] CRM systems are implemented in practice with a focus on the strategic application, in which analytical CRM systems are used to support customer knowledge acquisition and system development. A study describes CRM in higher education setting [5]. Whereas, it has been investigated that CRM and business-to-business (B2B) interactions are essential elements for the success of modern business. Although these are two different modules but they have many similarities. The integration of CRM and B2B can benefit all related prices in business processes including sales, marketing, customer service, and information support [6]. CRM applications can be categorized into operational and analytical [7]. Thorough analysis of the concepts of business intelligence (BI), knowledge management (KM) and analytical CRM (aCRM) and established a framework for integrating all the three to each other by using data mining (DM) techniques which would help in improving enterprise decision making [8]. It was examined and developed a better understanding of triangle relationship between quality, CRM and customer loyalty (CL) which might lead to companies competitiveness (CC) [9]. Whereas, data mining is performed using three selected software to databases related to customer survey, marketing campaign data, and web site usage. The three selected software are Poly Analyst of Mega pure Intelligence, Inc., SPSS Clementine, and Click Tracks by Web Analytics [10]. It was developed more comprehensive understanding of data mining by the examining the application of this technology in the market place. In addition to explore the technology that arises from the use of these applications, some of the social concerns that are often ignored are discussed [11]. Concern for business and emergence of the digital economy has made the problem even more acute. Companies have initiated programs to handle churn and customers profitability. The research proposed analysis framework able to prefigure the possible impact induced by the ongoing data mining enhancements on churn management and on the decision making process [12]. Top-management championship practices, employee information technology skills, and CRM knowledge are identified and examined as key building blocks toward strategic utilization. The empirical test of the conceptual model is based on a mail survey of North American firms that have adopted information technology-based CRM systems. The results based on random effects model show that strategic utilization of CRM technology leads to higher performance when there is an emphasis on using it to manage business-to-business rather than business-to-consumer relationships [13]. It was identified that success depends on skilled information technologists, among whom are statisticians are important. This research focuses on some of the contributions that statisticians are making to change the business world, especially through the development and application of data mining methods. This is very large area and the topics covered are chosen to avoid overlap with papers in this special issue, as well as to respect the limitations of expertise [14]. A twofold study showed that first in order to understand mobile customer relationship

management (mCRM) and data mining application in the mCRM, this paper aims to present a conceptualization of mCRM in respect of data mining. Secondly the paper also aims to develop the empirically grounded framework of the mCRM from data mining perspective.

The data mining for CRM has direct impact of knowledge management, technology, direct marketing and innovation but not been reported yet. Therefore, the objective of this research to find out the relationship of data mining on CRM with the independent variables such as knowledge management, technology, direct marketing and innovation. Accordingly, several research hypotheses are established to examine the relationship between four dimensions of CRM with data mining [15].

This research paper is organized as follows: Section 2 describes the theoretical framework and hypothesis, section 3 illustrates data analysis and results, Section 4 focus on the discussion about the results and at last the conclusion are drawn in Section 5.

2. Theoretical Framework and Hypothesis

2.1. Theoretical Framework

The theoretical framework contains four independent variables and one dependant variable namely knowledge management, innovation, direct marketing, technology and data mining for CRM respectively. We consider that there is direct relationship between knowledge management, innovation, direct marketing and technology with data mining for CRM.

2.1.1. Knowledge Management

Knowledge management in customer relationship management is independent variable which plays vital role in making organization focus on its retention of customers and by this they understand the needs and demands of customers [4]. It is hypothesized that knowledge management plays an important positive role in customer relationship management, implies it has a positive relationship.

2.1.2. Innovation

Innovation [3] is independent variable that is effected by Customer Relationship Management. Generally firms are able to increase their innovation capacity by ad hoc CRM. Like supplier integrated programs to form complete innovative programs. It is hypothesized that innovation plays an important positive role in customer relationship management, implies it has a positive relationship

2.1.3. Direct Marketing

[1] has developed a framework for prescribing an optimal mailing policy for a direct marketer. The model allowed for forward-looking behaviour for both firm and customer. It is hypothesized that direct marketing plays an important positive role in customer relationship management, implies it has a positive relationship.

2.1.4. Technology

[16] Depicts that it is very important and involved independent variable in customer relationship management because in current scenario we cannot forget the impacts of technology on different areas like HRM, OD, CRM etc. It is hypothesized that technology plays an important positive role in customer relationship management, implies it has a positive relationship.

2.1.5. Data Mining

Data mining has a positive relationship with data base security [11, 16] [17]. It is hypothesized that data mining has a positive relationship with customer relationship management

2.2. Theoretical Framework

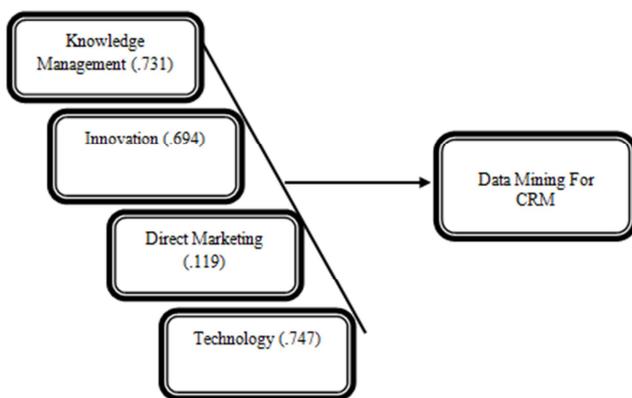


Fig 1. Theoretical Framework

2.3. Hypothesis

H₁: There is a relationship b/w Data mining and Knowledge Management

H₂: There is a relationship b/w Data mining and Innovation

H₃: There is a relationship b/w Data mining and Direct Marketing

H₄: There is a relationship b/w Data mining and technology.

3. Data Analysis and Results

The data for the research was collected through self-administered questionnaires from the students of Pakistani top universities. The questionnaire contains 15 questions in total, on the one independent and four dependent variables. The data was collected by randomly distributing the 175 questionnaires among the students and then collecting the filled questionnaires. The sample for this research data gathering was 175 randomly selected students. The purpose was to take a broad view of the result of the research to the huge population who have provided there data to any kind of record. The sampling method utilized for the reason of sampling was simple random sampling as we have randomly selected the students who filled the self-administered questionnaires.

3.1. Reliability Analysis

Cronbach's alpha is the most common measure of internal consistency ("reliability"). It is most commonly used when you have multiple Likert questions in a survey/questionnaire that form a scale and you wish to determine if the scale is reliable. In other words reliability is that other things being equal, a person should get the same score on a questionnaire if they complete it at two different points in the time. We can also say reliability to be that two persons, who are the same in terms of construct being measured, should get the same score. The main idea is that set of items should produce results consistent with the overall questionnaire. In a reliable scale the Cronbach's alpha should not be less than .65. Since the Cronbach's alpha for all our scales are greater than .65 so we can say that our scale is reliable. For knowledge management scale the Cronbach's alpha is .758 and the number of item is 5. since the Cronbach's alpha for data mining scale is above .65 so we can conclude it to be reliable. For data mining scale deletion of any item will not bring a substantial increase in the Cronbach's alpha. The items for data mining scale are correlated with overall scale since the value for all items is above .3. For Innovation scale the Cronbach's alpha is .841 and the number of item are 3, since the Cronbach's alpha for innovation scale is above .65 so we can conclude it to be reliable. Innovation scale deletion of item 3 will increase the Cronbach's alpha to .841 but this increase is not substantial increase in the Cronbach's alpha. The items for database security scale are correlated with overall scale since the value for all items is above .115. For Direct Marketing the Cronbach's alpha is .659 and the number of item is 4, since the Cronbach's alpha for privacy scale is above .65 so we can conclude it to be reliable. For Technology the Cronbach's alpha is .680 and the number of item is 3, since the Cronbach's alpha for privacy scale is above .65 so we can conclude it to be reliable.

Table 1. Reliability analysis

Variables	α
Knowledge Management	.758
Innovation	.841
Direct Marketing	.659
Technology	.680
Data Mining	.726

3.2. Descriptive Analysis

The maximum number of respondents are of age 22-25 because most of respondents are from BBA 5th -8th semester and in these semester most of students are about that age then 17-21 years of age groups and most of female respondents are about that age. The majority number of respondents who filled the questionnaires was males. The reason for this majority of male respondent was that male students were easier to approach than female students. Despite of the most of male respondent we can generalize our finding to the get their point of view about effective customer relationship management and usually we focused on management science department students because they are having customer relationship management knowledge.

Table 2. Descriptive analysis

Age	Respondents
17-21	69
22-25	93
26-30	13
Total	175

Table 3. Respondents

Gender	Respondents
Male	124
Female	51
Total	175

The descriptive analysis describes the overall characteristics of the data collected. The descriptive analysis table above described the characteristics of data collected on 15 items. The descriptive analysis reveal the number of respondent for each item, the range of the data collected, the minimum and the maximum value of the data collected, the mean of the data, the standard deviation of the data, the variance and Skewness of the data collected. The number of respondent for all items is 175, the range is 4 and the minimum value and maximum value for all is 1 and 5 respectively. The lowest mean is for item 2 that is 3.71 and the highest mean is for item 4 that is 4.04. This revealed that on item 2 respondent least positively inclined toward agreeing and on item 4 the respondent were most positively inclined to agreeing. For all items except item 2 and item 4 respondents were more inclined toward agreeing with the items. The average score for all items except item 2 and item 4 is above 3 showing respondent response to be on the right side of the neutral category and inclined toward agreeing with the items

Table 4. Descriptive statistics

	Mean	Std. Deviation	Variance	Skewness	
	Statistic	Statistic	Statistic	Statistic	Std. Error
Q1	3.9657	.98201	.964	-1.036	.184
Q2	3.7143	1.10826	1.228	-.695	.184
Q3	3.7886	.99186	.984	-.636	.184
Q4	4.0457	.92734	.860	-1.097	.184
Q5	3.8571	1.03232	1.066	-1.105	.184
Q6	3.9657	.98201	.964	-1.036	.184
Q7	3.7886	.99186	.984	-.636	.184
Q8	3.8571	1.03232	1.066	-1.105	.184
Q9	3.9657	.98201	.964	-1.036	.184
Q10	3.9657	.98201	.964	-1.036	.184
Q11	3.8571	1.03232	1.066	-1.105	.184
Q12	3.9714	.93728	.878	-1.044	.184
Q13	3.7429	1.09184	1.192	-.920	.184
Q14	3.7657	.92660	.859	-.787	.184
Q15	3.8343	1.00629	1.013	-.791	.184

3.3. Range

The perception of maximum people about all variables (knowledge management, technology, innovation, direct marketing) is about 4.8 and 4.9 so when we round these figure this will be “5” that means strongly agree and minimum its 2.9 that means “3” means neutral.

Table 5. Range

Variable's	Range
Knowledge Management	4.864 → 2.856
Technology	4.86 → 2.86
Innovation	4.95 → 2.97
Direct Marketing	4.80 → 2.83

3.4. Correlation Analysis

A correlation tells how and to what extent two variables are linearly related. The correlation is used when two variables that are quantitative in nature and are measured on a scale that is approximately interval scale. Correlation shows whether there is a relationship between the two variables and whether one variable is caused due to the other. The significance is the p-value that tells us the probability we would expect our result given the null hypothesis is true. If our p-value is less than alpha - .5 – we would reject the null hypothesis and conclude that there is a relationship between the two variables under study. In the table below correlation analysis for data mining and customer knowledge management, innovation, technology & direct marketing is shown. The correlation between the data mining and other variable is 1, which is strong positive correlation. The p-value is .000, this consistent with the correlation. The p-value is less than alpha therefore it is statistically significant. Thus we reject the null hypothesis and conclude that there is a relationship between data mining and customer relationship management.

Table 6. Correlation

	DAM	KM	TEC	INN	DM
DAM	1.000	.731	.747	.694	.119
KM		1.000	.918	.858	.055
TEC			1.000	.909	.099
INN				1.000	.034
DM					1.000

3.5. Regression Analysis

A regression is typically used to predict a dependent variable with an independent variable. Regression analysis is closely related to correlation analysis. In this research study the researcher is interested in whether data mining was related to knowledge management- innovation – technology – direct marketing. Whether there is a correlation is answered by correlation analysis. In this research simple regression is applied on two models. The model summary gives the adjusted R². The R² should be above .30 than it should be involved in the relationship. The R² for our model 1 Data mining and Database security relationship is 0.575.

Table 7. Regression Analysis

R	R Square	Adjusted R Square	Std. Error of the Estimate
.758 ^a	.575	.565	.50213

3.6. ANOVA

The table below shows the ANOVA analysis for data mining and knowledge management, innovation, technology

and direct marketing impact. According to the last box the p-value is less than .002. Since the p-value is less than alpha so we can make decision on the null hypothesis and conclude that there is a relationship or not a relationship between data mining in Customer Relationship Management and Knowledge Management, Innovation, Technology and Direct Marketing. The F-value is greater than 12 so it supports our conclusion drawn by the p-value.

Table 8. ANOVA

	Sum of Squares	Df	Mean Square	F	Sig.
Regression	57.915	4	14.479	57.425	.000
Residual	42.863	170	.252		
Total	100.777	174			

3.7. Coefficients of Regression

In the equation $Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 - e$, Y is the dependent variable, a is a constant, those are the beta coefficient that indicate the degree of influence the corresponding independent variable has on variations in the dependent variable. The higher the value of b, the more influential the independent variable and the sign indicates the nature of relationship whether the variable move together or in opposite direction. The Xs are the independent variables. The t statistic column is not so important to consider but the sig. column is important to see. The significant level should be below 10 percent. In regression table Innovation significance value is .608 and that is the greater value amongst all and that can cause massive change as a variable in CRM and that value is insignificant.

Table 9. Coefficient of regression

Variables	Standardized Coefficients	t	Sig.
Data Mining	-	2.477	.014
KM	.289	2.263	.025
TEC	.418	2.621	.010
INN	.063	.514	.608
DM	.060	1.170	.243

4. Discussion

By Data mining technique we can grasp all necessary information about CRM and by relating them with knowledge management and found very strong relationship between those variables. It's the very effective driver. This study showed that in the presence of non-confidential attributes the problem of grasping knowledge about customers and integrate them with the innovation customer preferences are contingent with data mining. One way to control the customer's information and their choices of data mining is by providing the individual with more control of their data that they have provided to any database. User defined sensitivity factors can minimize these issues we can solve them by using technology and enable customization of products. This method allows producers and service providers to rate their current products and services according to a specified scale of data mining.

This method allows us to know current market trends owner to know which data is more sensitive and which is not. By knowing the relevant data about the customers and then owner can utilize it freely in automated data analysis. Thus by this organizations will be better able to maximize innovation and customers appeal towards product or a service. Efforts should be made to get only relevant information and be in direct contact with suppliers and also customers (wholesalers and retails). The solution to any marketing issue should not only be applicable but it should also be compatible with the ever changing technology. Timely identification of customers related issues caused due to technological advancement in data mining, customers preferences changing very rapidly and should dealt as they arise and derive their solution in a timely and concurrent manner. The use of historical archives for data mining purpose can reveal many customer related issues. One should be aware of the consequences of these issues. The choice of appropriate data analysis technique can have a great effect on the data mining for Customer Relationship Management that can be caused by it. Data mining is a new field and there exists a lot more to explore in this technology. As new methods of data mining are invented the problem contingent with them will also arise.

5. Conclusion

In this research tried to make effort to find out the relationship between customer relationship management and knowledge management, innovation, technology and direct marketing. The research study is about the challenges facing CRM (Customer Relationship Management). The objective of research is to find the relationship b/w knowledge management and the data mining to analyze target customers and know needs and the wants of customers or their information like their emails, contact number and addresses. To investigate personal selling relationship with customer relationship management, how they react, feel or perceived about the existing or the newly launched products. To find out the relationship between customer involvement and process innovation; customer involvement and administrative innovation; and long-term partnership and marketing innovation are not significant and technology-based CRM has positive effects on all five types of innovation. To search another challenge related to managing customer development pertains to the data used to model response. Companies nowadays record their transactions with each individual customer and store those in customer transaction databases. Rather than information on a sample of customers, companies in question have access to the entire population of customer.

The results of our research are in stripe with past researches done in this area. The correlation between the dependent variable – data mining and the independent variables – knowledge management, innovation, technology and direct marketing is positive and strong. The p-value for both the correlation analysis is also in conformity with our correlation value and thus concluding that there is a relationship between the data mining and knowledge management, innovation,

technology and direct marketing the R square values also confirm our above finding.

The results, based on random effects model, show that strategic utilization of CRM technology leads to higher performance when there is an emphasis on using it to manage business-to-business rather than business-to-consumer relationships, user expertise (but not buy-in) impacts CRM performance through strategic utilization, and top management championship practices, CRM knowledge, and employee IT skills impact strategic utilization through buy-in and expertise.

- Markets should get information about customers that helps in facilitation to customers and help them to provide service in better way.
- They should update their technology and implement in daily processes
- Innovation helps in growth of product line or product/service Quantum.
- By knowledge management it helps in pricing decisions, manufacturing highly customized products and design as according to customer preferences.
- Direct Marketing helps in making long-term and strong relationship with customers, retailers and whole sellers.

References

- [1] W. Kamakura, C. F. Mela, A. Ansari, A. Bodapati, P. Fader, R. Iyengar, et al., "Choice models and customer relationship management," *Marketing Letters*, vol. 16, pp. 279-291, 2005.
- [2] C. Bull, "Strategic issues in customer relationship management (CRM) implementation," *Business Process Management Journal*, vol. 9, pp. 592-602, 2003.
- [3] R.-J. Lin, R.-H. Chen, and K. Kuan-Shun Chiu, "Customer relationship management and innovation capability: an empirical study," *Industrial Management & Data Systems*, vol. 110, pp. 111-133, 2010.
- [4] M. Xu and J. Walton, "Gaining customer knowledge through analytical CRM," *Industrial management & data systems*, vol. 105, pp. 955-971, 2005.
- [5] E. D. Seeman and M. O'Hara, "Customer relationship management in higher education: Using information systems to improve the student-school relationship," *Campus-Wide Information Systems*, vol. 23, pp. 24-34, 2006.
- [6] Y. E. Zeng, H. J. Wen, and D. C. Yen, "Customer relationship management (CRM) in business-to-business (B2B) e-commerce," *Information Management & Computer Security*, vol. 11, pp. 39-44, 2003.
- [7] P. Raman, C. M. Wittmann, and N. A. Rauseo, "Leveraging CRM for sales: the role of organizational capabilities in successful CRM implementation," *Journal of Personal Selling & Sales Management*, vol. 26, pp. 39-53, 2006.
- [8] J. Ranjan and V. Bhatnagar, "Role of knowledge management and analytical CRM in business: data mining based framework," *The Learning Organization*, vol. 18, pp. 131-148, 2011.
- [9] R. C. Leventhal and M. Zineldin, "The royalty of loyalty: CRM, quality and retention," *Journal of consumer marketing*, vol. 23, pp. 430-437, 2006.
- [10] R. S. Segall and Q. Zhang, "Web mining technologies for customer and marketing surveys," *Kybernetes*, vol. 38, pp. 925-949, 2009.
- [11] A. Danna and O. H. Gandy Jr, "All that glitters is not gold: Digging beneath the surface of data mining," *Journal of Business Ethics*, vol. 40, pp. 373-386, 2002.
- [12] M. A. Lejeune, "Measuring the impact of data mining on churn management," *Internet Research*, vol. 11, pp. 375-387, 2001.
- [13] A. Saini, R. Grewal, and J. L. Johnson, "Putting market-facing technology to work: Organizational drivers of CRM performance," *Marketing Letters*, vol. 21, pp. 365-383, 2010.
- [14] D. L. Banks and Y. H. Said, "Data mining in electronic commerce," *Statistical Science*, pp. 234-246, 2006.
- [15] J. Ranjan and V. Bhatnagar, "A holistic framework for mCRM-data mining perspective," *Information Management & Computer Security*, vol. 17, pp. 151-165, 2009.
- [16] L. Y. Sin, A. C. Tse, and F. H. Yim, "CRM: conceptualization and scale development," *European Journal of Marketing*, vol. 39, pp. 1264-1290, 2005.
- [17] A. Payne and P. Frow, "A strategic framework for customer relationship management," *Journal of marketing*, vol. 69, pp. 167-176, 2005.