THE ROLE OF DATA MINING IN CUSTOMER RELATIONSHIP MANAGEMENT SYSTEMS

[1]EHSAN VASEGHI, [2]HASSAN ESKANDARI

[1,2]SHAHIDCHAMRAN University Iran Ahvaz
E-mail: vaseghi@live.com, hassan.eskandari@live.com

Abstract - In the past, analysts generally were tasked with extracting useful information from recorded data manually. But since this manual data analysis is very expensive and the data volume and integrity are increasing day by day moving from straight-line analysis to the automatic indirect ones, utilizing of Computer methods such as data mining is an inevitable option to cope with the complexity of the problem. Data mining refers to the extraction and analysis of large amount of data in order to extract data patterns, increase their core value and transfer data and knowledge. Nowadays, future companies, which prioritize their system in the terms of customer relationship, benefit from data mining approaches to recognize all customers and their needs which in turn, helps to have a better understating of their priorities and consequently to establish a strong relationship with them. Such a relationship with customers increases the competing capabilities of companies in the market. Data mining makes it possible to determine the value of customers, predict their future behavior and make informed decisions in this regard. This study was conducted in a systematic and exhaustive way to consider that to what extent companies have been using the data mining approaches in their business in the recent years. Results showed the tendency to use data mining techniques in customer relationship management systems, marketing and attracting customers and increasing customer loyalty has increased.

Keywords - Data Mining, Customer Orientation, Customer Relationship Management, Clustering

I. INTRODUCTION

The customer relationship changes with time, and whatever the business company and the customers know each other more, this relationship evolves and grows. The customer life cycle provides a good framework for employing data mining in customer relationship management. In the Data Entry, the customer life cycle tells us what information is available and in the output, the life cycle shows what is likely to be interesting and what decisions should be made. Data mining can predict the profitability of potential customers that can be turned into actual customers, and how long they will remain loyal to the company and how likely they will leave [1]. Some customers regularly visit a certain company while others change companies to gain benefit. In this case, companies can focus on customers that are more profitable [2].

Data mining is one of the recent advances in data management technologies which is described as a set of techniques that allow a person to move beyond ordinary data processing and extract specific information that is hidden among a mass of data. The motivation to expand data mining was largely due to the change in the world of commerce in 1990s. Data mining in the field of marketing, for example, is particularly helpful to alleviatethe unpredictability of the relationship between the clients' profiles and their buying pattern. [3]Data mining knowledge has flourished with the advancement of computer processors and the ability to store data retrieval and analysis, and not only has answered many questions and needs of researchers and managers of small and large businesses but the ordinary people has been benefited of it. Using a variety of sciences such as mathematical statistics and artificial intelligence, this knowledge will respond to the new needs, including predicting and discovering hidden patterns among a bunch of data. Considering the intense competition in the market and the variety of products and services offered to customers, customer behavior analysis is considered to be a very important factor for the survival of the firms.

One of the most effective applications of data mining is to examine the behavior of clients. Efficient customer relationship management is only achieved if the organization has sufficient understanding of customer preferences and requirements. This can lead to the development of profitable strategies and therefore companies are looking for a way to identify customers and understand their needs as much as possible. For this purpose, a strong tool is needed beyond the simple retrieval of the information to analyze the customer's behavior. In recent years, data collection technologies have been growing rapidly, and recently the issue of corporate advancement, in particular in the area of financial decision-making, has shifted data collection from acquiring the power to extract useful knowledge in the data. Although many organizations hold a huge amount of valuable data of their current customers, potential customers, suppliers and business partners, the inability of organizations to apply the appropriate techniques to analyze the data prevents them from discovering the valuable findings available in this massive data set. In such a situation, data mining takes an important role to solve the problem. In fact, customer relationship management and marketing can be considered as one of the first areas of data mining. Data mining tools answer questions that in the past, took a long time to be

addressed, if not impossible.
The main issue in this study is the evaluation of the role of data mining in customer relationship management systems.

II. RESEARCH LITERATURE

New information and communication technologies, as well as decision support technologies, can collect, store, evaluate, interpret, analyze, retrieve and disseminate information and knowledge to specific users. Data mining involves the use of advanced data analysis tools to discover valid, previously unknown patterns and relationships in a large data set [11]. Data mining is more than collecting and managing data. Indeed it includes analysis, prediction and discovery of knowledge in the database(Fig1).

Data mining can be done on quantitative, textual, or multimedia data. Its applications include:

- **Dependency rules**: Patterns in which the existence of an item implies the existence of another item.
- **Classification**: Assigning patterns to a small set of classes predefined by discovering some relationships between properties,
- **Clustering**: A group of customers or a set of patterns that have similar characteristics,
- **Prediction**: Discover patterns for a predictive logic about the future,
- **Path analysis or sequential patterns**: Patterns in which an event results in another occurrence.

Data mining is not a new technology, but its application has been growing significantly in various private and public sectors, and generally the industries such as banks, insurance, medicine and retailers have been using data mining to reduce costs, increase research and increase sales. [4]
Some of the data mining applications in real-world environments include:

Retail: It is a classic data mining application that can be found in:
- Determine Customer Purchase Patterns
- Market basket analysis.

Banking:
- Predicting fraud patterns through credit cards
- Diagnosis of fixed customers
- Determine the number of credit cards used by social groups.

Insurance:
- Case analysis
- Prognosis of new insurance policies by customers

Medical services:
- Determine the type of behavior with patients and predict the success rate of surgery
- Determine the success rate of treatment methods in dealing with severe diseases

The process of discovery consists of five steps:
✓ Data storage
✓ Data selection
✓ Data conversion
✓ Explore the data
✓ Interpreting the result

As we can see, data mining itself is the fourth step of the discovery process which plays an important role in the discovery of knowledge from the data.

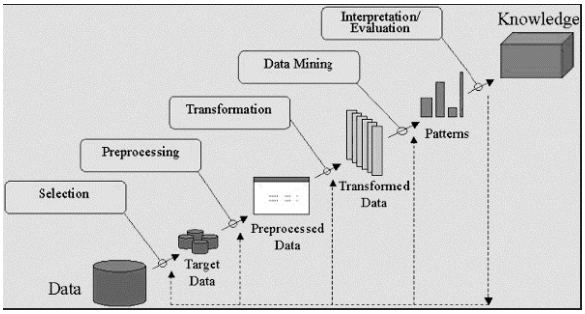


Figure I

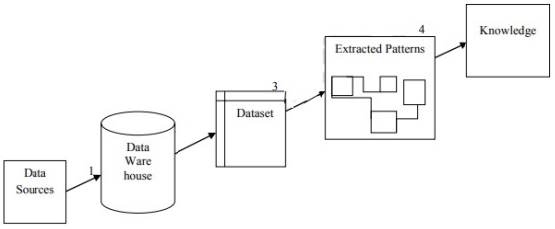


Figure II : The Process of data mining

Data Mining involves the following steps:

1) **Data Cleansing:** It means to clean the available datawhich means removing noises and inconsistencies from data and store it in a data warehouse. The data that is stored in the data warehouse should be cleaned data.
2) **Selection and Transformation**: Data selection means to select only required data i.e. data which is needed to solve a certain problem. Data transformation means to transform the data into various forms which are required for mining.
3) **Data Mining:** In this step, a number of methods are applied to extract or generate data patterns from the available and pre-existing dataset. It can be done using various techniques and algorithms such as clustering, classification, regression analysis, etc. Some of them are explained in the next section.
4) **Post-processing and Visualization**: At last, post-processing of extracted data is done which can be resulted in desired or undesired

outcomes and regardless of the degree of desirability the outcomes are visualized into an understandable format for users.

The extracted data in the definition of data mining implicitly refers to the information that can be derived from the results that are not normally tangible. Under this definition, the availability oflarge databases or a large amount of the processed data is of utmost importance

Data mining tools rely heavily on information and knowledge-based decision-making processes, and today's modern business activities rely heavily on it.

Nowadays, taking some measures such as making accurate and timely decisions, use of appropriate strategies for categorizing and finally finding user-friendly and efficient user data from a wealth of data is of primitive necessities to survive in the market. These measures need the access to the relevant information to be taken. However drawbacks such as wide variety of contacts, customers, the diversity and complexity of services and business environments make it difficult to have access to this relevant information. Data mining is a response to this need of organizations and institutions. The greater volume of the data and the more complicated relationship between them, the more difficult access to data. So the role of the data mining as one of the methods of discovery becomes clearer.

Today, the use of traditional methods of collecting and analyzing data is not appropriate due to the time and huge money costs and instead, new data analysis methods such as data mining can be a proper alternative. Competitiveness of global and regional markets requires institutions to create competitive advantages in providing service delivery, increasing satisfaction and attracting more customers in order to develop the organization.

According to the some theorists of information and statistics technology, data mining is the basis of simulation. The term data mining has been used by statisticians, data analysts and the information management community.

Scientific data mining is an integral part of the relationship between statistics science, computer science, artificial intelligence, modeling, machine learning and visual representation of data. This activity is a complex process to identify patterns and models that are accurate, new, potentially useful in a wide range of data and understandable to humans. The statistical units of the organizations are highly recommended to use this science seriously to get the relationships between intelligent data and other parts.[6]

A user of knowledge discovery systems should have a high degree of understanding of the scope of the data in order to select a right data subcategory, the appropriate category of patterns, and a good standard for attractive patterns. Therefore, database discovery systems must have interactive tools, not automated analysis systems. Therefore, the process of discovery of knowledge from the database should include the following steps:

1) Understanding the realm
2) Preparing the dataset
3) Pattern Discovery (Data mining)
4) Processing after pattern discovery
5) Using the results.

III. CRM DEFINITION

The CRM strategy has many definitions. Some writers consider it as a more complete form of relational marketing, and define it as creating, developing, and strengthening individual relationships with customers in order to maximize the total value of their life cycle. Another perspective for CRM is seen as a technology defined as Creating and maintaining a personal relationship with profitable customers through the proper use of information and communication technologies[5]. Another perspective is that customer relationship management is a leader or philosophy. In the most complete way, CRM is a strategy or philosophy, not a technology specially designed to optimize profitability, revenue, and customer satisfaction.

CRM can also be seen as a management approach that enables organizations to identify and attract profitable customers by managing relationships with them and leads to increased customer retention. But one of the best definitions is the overall process of creating and maintaining beneficial communication with the customer by delivering the value of the key customers and satisfying them[5]. Another definition states that customer relationship management is a comprehensive business and marketing strategy that integrates technology, processes, and all business activities around a customer. [7]

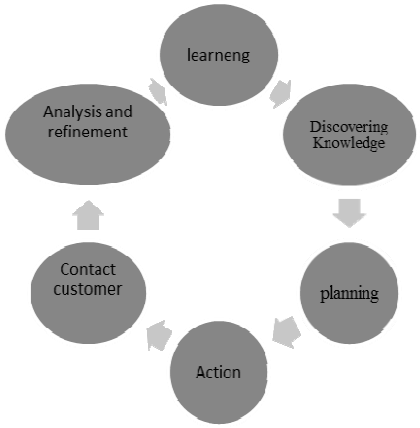


Figure III CRM Cycle (Tam 2002)

In [8]using data mining, researchers clustered and categorized customers of a bank to identify risky loans.

In another article in 2003 [9], considering the share of profits, the potential profit and customer profitability a

definition for a model was proposed and the customers were divided accordingly.

IV. DATA MINING PHASES IN CUSTOMER RELATIONSHIP MANAGEMENT

Using a methodology is common in many customer relationship management projects. One of the most prominent methodologies in this area is the Crisp methodology which follows the steps below to perform a data mining project in the field of customer relationship management.

1) **Understanding the business environment**: At this stage, the business is recognized. The status quo is measured and the questions that the data mining should answer.
2) **Recognition of data**: In this phase, the raw data in the business environment is collected and reviewed. At this point, it is decided that whether the modeling takes place on the whole data or only the sample data. Identifying appropriate data sources and collecting them is an important key factor.
3) **Selecting and preprocessing data**: The purpose of this step is to extract relevant data for analysis from a larger data warehouse. This phase clears data and prepares them to ensure that they deliver the right results. Deleting lost data ensures that values of the same meaning are included in this set (For example the uniformity of monetary units). Normalization to ensure that the results are not converged to high-value placeholders is the task of this step.
4) **Modeling**: Modeling is done using the existing data mining techniques. Types of data mining techniques are categorized based on the outputs they generate. The cycle continues until an acceptable model is presented. The process of selecting a data mining technique depends on the nature of the data and the business purpose.
5) **Evaluation of results**: After modeling, the model has to be evaluated based on the criteria set for the model and measures are taken to modify or improve the model if necessary.
6) **Improvement**: The model is developed and improved on the basis of exerting corrective actions on earlier stages. Then, the results are re-evaluated and the recovery process continues to reach the desired response.

V. DATA MINING APPLICATIONS IN CUSTOMER RELATIONSHIP MANAGEMENT

In the area of customer relationship management and marketing, data mining can be considered as a process of the discovery of profitable knowledge derived from organizational data and continuous use of them. Data mining can also be used to properly orientate the decision-making process and predict the impact of the decisions. For example, data mining expedites the organization's response to customer requests by classifying them into groups with different characteristics. It can also calculate the probability of current customer migration to the other companies. Data mining is especially important for companies in which the cost of switching providers for their customers is very low, so the organization should promote customer retention programs. By knowing who, why and even when they want to leave the organization, a maintenance plan for the valued customers can be planned.

By examining customer demand and behavior, it is possible to extract rules as a sales plan, new strategies for the marketing department, that will have a great deal of sales forecast. The activities of an organization should include delivery of the correct product to the target customer through the right channel at the right time and at the right price. Usually a combination of data mining models is needed to support and predict the impact of customer relationship management strategies. After a thorough analysis and evaluation, the organization's actions may include responding to customer specific requests, anticipating increased sales, some new marketing plans, new advertising strategies, and guidelines for employees in different sectors[10].

Faced with real issues, what is visible is only raw data. Data mining determines of an existing set of methods which one can be found in business-friendly strategies and eventually achieve one of the key goals.

Data mining techniques represent multiple algorithms that each one is a set of data mining techniques described in the previous section. Each method pursues a specific goal, for example, clustering of data into distinct clusters. The classification after the model training phase addresses the prediction of the unlabeled data category. All businesses are interested in anticipating the behavior of their customers through the knowledge gained through the data mining.

The goal of Customer Relationship Management is to create a better understanding and recognize the needs of customers to provide a better service and thus to keep the profitable customers. Here, data mining is a valuable tool for discovering the needs and identifying customers from a massive amount of data. Data mining will meet the information needs by providing methods such as prediction, categorization and clustering. The use of data mining to detect and communicate with the customer has been frequent in many applications.[11]

Clustering in CRM

Clustering is a process that helps us identify similar data sets and understand both the similarities and differences within the data. Data sets that have similar traits can be used for conversion rate increases. For example, if the buying behavior of one group of

customers is similar to that of another group, they can both be targeted with similar services or products.[12]

Regression in CRM

Regression analysis is one of the advanced data mining techniques in CRM. It tries to find the dependency between different data items and shows which variables are affected by other variables. Unlike correlation, however, regression only shows those variables that are affected by others, and does not indicate which variables can affect others. This technique is used to determine the customer satisfaction levels and how- and to what degree- they affect customer loyalty.[13]

Data mining techniques in CRM help businesses find and select relevant information that may be used to analyze and predict customer's behavior. The more data available in the database, the better models can be created by data mining techniques which consequently will result in more business values[14].

VI. RESEARCH METHODOLOGY

The method used for this study was a descriptive-analytic and systematic review of previous studies carried out on the subject. We have used Google Scholar, Scopus, and Thomson Reuters to collect data. We looked for keywords like data mining and customer orientation dated between 2008 and 2018.

FINDINGS AND CONCLUSIONS

The results of the study showed that in the last decade, the tendency to use data mining techniques (clustering, genetic algorithm, regression, Markov chain, decision tree, survival analysis, neural networks, etc.) in managing communication systems with customers, process improvement, marketing and attracting customers and increasing customer loyalty and satisfaction has increased significantly. Companies that use this approach are more successful in meeting their goals. In this way, they will invest in identifying and keeping high-valued customers, find out customers who are less likely to profit and make decisions on how they will deal with them. Using data mining techniques will empower decision makers to provide information about the future of their systems and develop targeted plans for their company's perspective/outlook [15].They will also be able to take standing on the perceived position of the business environment rather than the impact of the business, and to rank themselves in this competitive business market.

RECOMMENDATIONS

- Increasing managerial information about data mining technology and its benefits
- The existence of integrated information infrastructures and normal databases informing programmers' awareness of the company's requirements about some of the information necessary to implement new technologies.
- Complete registration of all information in the database fields
- Increasing managers' awareness of the profitability of marketing
- Increasing customer surveys and respecting customer needs.

REFERENCES

[1] Gharavi, E. (2012). Customer churn management Using customer lifetime value. (Master Dissertation), Khaje Nasir University of Technology. [Persian]

[2] Farooqi, Rashid, Raza, Khalid 2011, "A Comprehensive Study of CRM through Data Mining Techniques.

[3] Abdou, H., & Pointon, J, (2008) "Neutral nets versus conventional techniques in credit scoring inEgyptian banking", Expert System with Application, Vol, xxx, pp.1

[4] Chen, M- Ch., Huang, Sh- H. (2003) "Credit scoring and rejected instances reassigning through evolutionary computation techniques", Expert Systems with Applications, 24[4], pp. 433- 441.

[5] Guillén, M., Nielsen, J. P., Scheike, T. H., & Pérez-Marín, A. M. (2012). Time-varying effects in the analysis of customer loyalty: A case study in insurance. Expert Systems with Applications, 39[3], 3551- 3558.

[6] Aarushi, Rainu Nandal, APPLICATION OF DATA MINING IN CUSTOMER RELATIONSHIP MANAGEMENT: A REVIEW, January 2018, Volume 5, Issue 1 JETIR (ISSN-2349-5162)

[7] Jussi Ahola, Esa Rinta-Runsala, (2002) "Data Mining Case Studies in Customer Profiling", VTT information Technology.

[8] Kim, Y. S., & Sohn, S. Y. (2004). Managing loan customers using misclassification patterns of credit scoring model. Expert Systems with Applications, 26, pp. 567– 573.

[9] Kantardzic, Mehmed, (2003), "Data Mining: Concepts, models, methods, and algorithms", wiley-inter science.

[10] Olson, D. &, Shi. Y., "Introduction to Business Data Manning. Singapore": MCGrow Hill Education, 2007.

[11] Zhang, Y., & Bhattacharyya,S., "Genetic Programing in classifying largescale data:an ensemble method".Information Sciences,vol 163, pp.2,6,2009.

[12] Tarakh Mohammad Jafar, Sharifian Kobra. (1386). Application of data mining in improving customer relationship management. Industrial Management Studies, 6 (17), 153-181.

[13] Wen-Yu Chiang,Applying data mining for online CRM marketing strategy: an empirical case of the coffee shop industry in Taiwan, British Food Journal, https://doi.org/10.1108/BFJ-02-2017-0075. Eastern Michigan University at 15:12 29 January 2018.

[14] Shaghayegh Abolmakarem, Customer Behavior Mining Framework (CBMF) using clustering and classification techniques, Journal of Industrial Engineering International, 11 August 2018.

[15] Dr. Miriam McGaugh, Clinical Assistant Professor of Marketing, Data Mining and Customer Relationship Management Applications, Marketing 5963, Spring 2018. MKTG 5963 – Data Mining and CRM.

★★★