

T & H Article: Data Mining for Customer Relationship Management

Nurharyati

*Fontys University of Applied Sciences
Information and Communication Technology*

ME_2_1@YAHOO.COM

About the author



My name's Nurharyati. I was born in Denpasar, Bali – Indonesia. Bali is the most beautiful island in the world, it has a lot of art, culture, and also tourist destinations. In 2006, I moved to Surabaya – East Java, to study Informatics Engineering at Petra Christian University. I have studied at Petra Christian University for 3 years and I decided to finish my study here. Now, I am taking Final Year program at Fontys University of Applied Science on ICT major. It is a good opportunity for me to get experience and learn how to work with people from different culture. I hope that I can graduate next year and find a good job here.

Abstract

In a company, it is important to have a good relationship with their customers. The relationship can be built and maintained by understanding customer's needs and wants which can be found from customer's activity. Along with development of technology, data mining application is required and used by a lot of companies to recognize customer behavior by collecting customer data and activity data. By using data mining, complexity of customer relationship can be handled, although there is not enough time, there are a lot of customers, a lot of products, and also a lot of competitors. Customer Relationship Management (CRM) helps companies to improve their interaction and relationship with their customers which can improve their profits and the customer satisfaction.

Keywords: data mining, customer relationship management

Introduction

Nowadays, a lot of data is collected. The amount of data increases constantly. By analyzing those data, we can discover new knowledge which is useful for business processes. Manual methods cannot analyze enormous amount of data that already exist in a company. Researchers found an opportunity to build a new technology to replace the weakness of manual methods, which is Data Mining. Data Mining has been implemented by many companies to solve their business problem. One of the implementations of data mining in industry is Customer Relationship Management (CRM).

Data Mining

Data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both [1]. So, data mining is defined as a process of discovering previously unknown and valuable information from a collection of data (data warehouse). It allows users to analyze enormous data amounts faster than manual method to solve business problem. Data mining explores databases to find undiscovered pattern, find information which is forgotten by businessmen.

Data mining is a step of Knowledge Discovery in Databases (KDD) Process: [2]

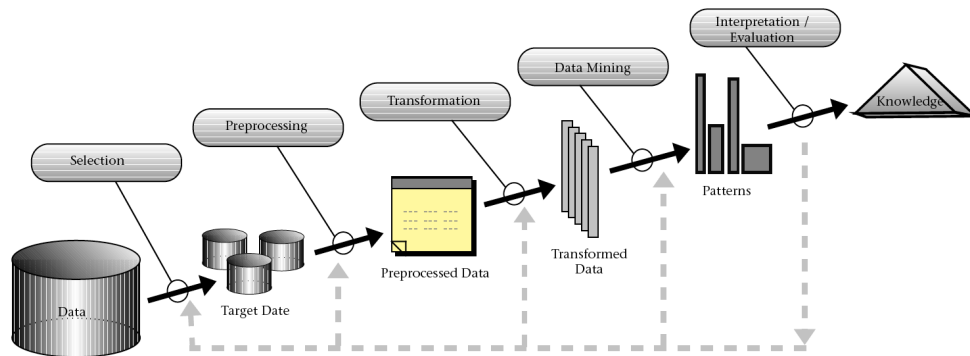


Figure1. Steps of KDD Process

To do Data Mining, we have to do some steps before it. First, we have to have a Data Warehouse or Data Mart. Then, we do Data Selection, Data Preprocessing, and Data Transformation. Data Mining is sometimes referred to as KDD, and Data Mining and KDD tend to be used as synonyms.

Data Mining has abilities as the following:

- Optimization of trend prediction and business behavior. Data Mining optimizes process of finding information and predicts in a huge database. Questions which are related to this prediction can be answered directly from the available data. Examples of prediction are marketing target, bankruptcy forecasting, and other kind of losses.
- Optimization of pattern finding which were previously unknown. For example, analysis of sales data to identify products that seems unrelated but oftentimes bought together.

Data mining commonly involves four classes of task: [2]

- Classification - Arranges the data into predefined groups. For example an email program might attempt to classify an email as legitimate or spam. Common algorithms include Nearest neighbor, Naive Bayes classifier and Neural network.
- Clustering - Is like classification but the groups are not predefined, so the algorithm will try to group similar items together.
- Regression - Attempts to find a function which models the data with the least error. A common method is to use Genetic Programming.
- Association rule learning - Searches for relationships between variables. For example a supermarket might gather data on customer purchasing habits. Using association rule learning, the supermarket can determine which products are frequently bought together and use this information for marketing purposes. This is sometimes referred to as "market basket analysis".

Customer Relationship Management (CRM)

Customer relationship management (CRM) is a method that companies use to interact with customers [3]. It manages the interaction between a company and its customer, and also to understand consumer behavior. CRM helps businesses use technology and human resources to gain insight into the behavior of customers and the value of those customers [4].

The advantages of CRM: [4]

- Provide better customer service
- Increase customer revenues
- Discover new customers
- Cross sell/Up Sell products more effectively
- Help sales staff close deals faster
- Make call centers more efficient
- Simplify marketing and sales processes

Acquiring, building, and retaining customer are the top priorities. The quality of customer relationship provides many companies to become more competitive than others. Many companies are realizing that to improve their position and global marketplace requires closer relationship with their customers. Improved customer relationships can improve profitability in three ways: by attracting more suitable customers, by generating profit through cross-selling and up-selling activities, and by extending profit through customer retention.

Data Mining in CRM Process

The most important implementation of Data Mining in industry is CRM. According to the polling made by KD Nugget on 107 companies, in 2008 the use of Data Mining in CRM took the first place in industry. The result of the polling in December 2008 is present on Figure 2 [5].

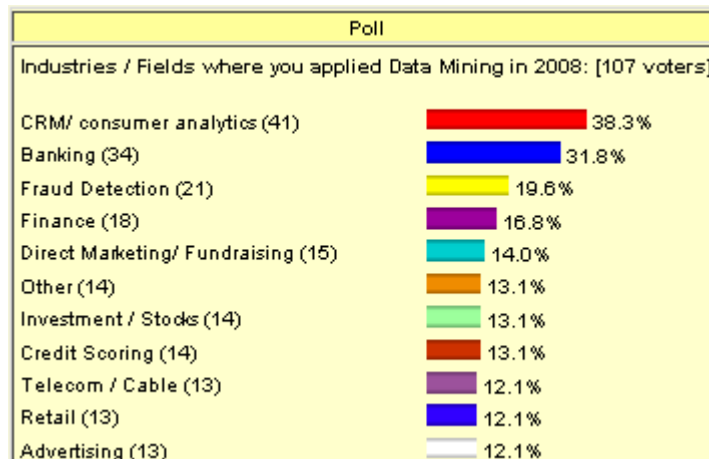


Figure2. Result of 2008 polling of Data Mining application in Industry

Data Mining can build a model of customer needs and behavior by using statistical and machine learning technique. Therefore, data mining applications can identify which market contains customers in high profit potential, by finding patterns among different variables that serves prediction about purchasing behavior.

Data Mining has many techniques to discover new knowledge from available data. There are some of data mining techniques that associated with CRM task. It is presented by Table 1 [7].

Data mining technique	CRM operation
Association rules	Information from customer-purchase histories is used to formulate probabilistic rules for subsequent purchases.
Decision trees	Automatically constructed from data, these yield a sequence of simple rules; good for identifying important predictor variables, nonlinear relationships, and interactions among variables
Genetic algorithms	Use procedures modeled on evolutionary biology to solve prediction and classification problems or develop sets of decision rules.
Neural networks	Applications that mimic the processes of the human brain; capable of learning from examples (large training sets of data) to discover patterns in data
Query tools	Provide summary measures such as counts, totals, and averages.
Regression-type models	Ordinary least-squares regression, logistic regression, discriminant analysis; used mostly for confirmation of models built by "machine learning" techniques
Visualization tools	Histograms, box plots, scatter diagrams; useful for condensing large data sets

Table 1. Association between data mining technique and CRM operation

1.1. Customer Life Cycle

Customer Life Cycle is the stages in the relationship between a customer and a business [6]. Data Mining helps companies to build good relationship by identifying customer needs and behavior throughout customer life cycle. Customer Life Cycle has three stages [9]:

- Acquiring new customers
- Increasing the value of existing customers
- Retaining good customers

There are some key stages of customer life cycle that can be gained by Data mining [6]. Key stages in the customer lifecycle:

- Prospects: people who are not yet customers but are in the target market
- Responders: prospects who show an interest in a product or service
- Active Customers: people who are currently using the product or service
- Former Customers: may be "bad" customers who did not pay their bills or who incurred high costs

Throughout Customer Life Cycle, Data Mining helps to determine the behavior surrounding a particular life cycle and also helps to find other people in similar stages and determine which customers are following similar behavior patterns.

1.2. Applying Data Mining to CRM

There are some steps for building good model of CRM system: [6]

- Define business problem
First, we should define the problem and the goal to build a model. It's because each CRM application has more than one business objective, so we need to build the appropriate model.
- Build marketing database
The second step is we need to build marketing database because operational database and corporate data warehouse will not often contain the data that we need.

- Explore data
Before we can build good predictive model, we must understand our own data by exploring the data which can be done by looking at distribution of data and gathering numerical summary.
- Prepare data for modeling
It is the final data preparation step before building model. First, we select the variables on which to build the model. Next is to construct new predictors from the row data, decide to select a subset of data on which to build models. Last, we need to transform variables in accordance with the requirement of the algorithm we choose for building the model.
- Build model
We need to explore alternative models to find the one which is most useful in solving our business problem.
- Evaluate model
After we build the model, we have to evaluate whether the model is appropriate enough to represent our problem. If it's not appropriate enough, we can loop back to previous steps to fix it.
- Develop model and result
Latest, the good model has been built, we can develop and applying this model.

Recently, marketers have added a new class of software to their targeting arsenal. Data mining applications automate the process of searching the mountains of data to find patterns that are good predictors of purchasing behaviors. After mining the data, marketers must feed the results into campaign management software that, as the name implies, manages the campaign directed at the defined market segments [8]. It is represent on Figure 3 [6].

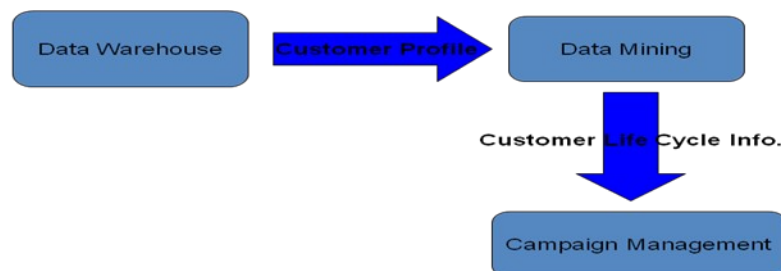


Figure3. Data flow that delivers information to Campaign Management

Example of CRM Benefit

The following is the example of CRM benefit to increase profit. At the beginning, companies will invest high cost for implementing CRM (software, computer, and people resources). But, after they use it for their business processes, it will increase their profit and got refund for the investment. This following example is about acquiring new customer with data mining where a company must invest to increase its profit is like invest for implementing CRM[9].

For example, a bank annually conducts many campaigns which offer one million people the opportunity for credit card. But, the portion of people who become credit card customers is only about 1% per campaign.

The first step is getting people fill out the application for the credit card. Then, the bank must decide whether the applicant is a good risk and accept them as a customer. While 6% of them respond with an application, only 16% of them are suitable credit risk. For a net, only 1% of all the mailing list becomes a customer.

The cost of mailing the solicitations is about \$1 per piece, totally \$1,000,000 per campaign. Over the next couple of years, the customers will generate \$1,250,000 profit for the bank, for a net return from the mailing of \$250,000. Data Mining can improve this return.

First, the bank did a test mailing of 50,000 and carefully analyzed the results, building a predictive model of who would respond and a credit scoring model. Then combining these two models to find the people who most likely to respond to the offer and have good risks.

The model was applied to the remaining 950,000 people in the mailing list which 700,000 people selected for the mailing. From the 750,000 pieces mailed overall, 9,000 acceptable application for credit card were received. It means that the response rate increases 20% (from 1% to 1.2%). The following table summarizes the result[9]:

	Old	New	Difference
Number of pieces mailed	1,000,000	750,000	250,000
Cost of mailing	\$1,000,000	\$750,000	\$250,000
Number of responses	10,000	9,000	1,000
Gross profit	\$1,250,000	\$1,125,000	\$125,000
Net profit	\$250,000	\$375,000	\$125,000
Cost of models	0	\$40,000	(\$40,000)
Final Profit	\$250,000	\$335,000	\$85,000

Table2. Differences before and after building customer model.

Even the bank include \$40,000 cost for building customer model (software, computer and people resources), the net profit increased \$85,000. It means that the high cost investment for implementing data mining (CRM) will be very useful to increase profit.

Conclusion

Data mining which use a large amount of available data helps businessmen to maintain and develop their business. By applying Data Mining application for CRM, a company can get a lot of benefit from that application. A company can understand which customers are more profitable than others. It also helps companies to keep the relationship with their customer to gain more profit.

Acknowledgements

I would like to thank to Mr. Geert Dirks and Mr. Peter Boots who have provided this template for writing this article. I also thank to my friends who have supported me during making this article.

References

<http://www.anderson.ucla.edu/faculty/jason.frand/teacher/technologies/palace/datamining.htm>, accessed: on 20 October 2009.

<http://www.kdnuggets.com/gpspubs/aimag-kdd-overview-1996-Fayyad.pdf>, accessed: on 20 October 2009.

http://en.wikipedia.org/wiki/Customer_relationship_management, accessed: on 23 October 2009.

<http://crmtutorial.com/CRM/CRM.aspx>, accessed: on 23 October 2009.

<http://www.kdnuggets.com/polls/2008/data-mining-applications.htm>, accessed: on 23 October 2009.

www.data-miners.com/companion/Chapter14-Example2.ppt, accessed: on 23 October 2009.

http://www.nbu.gov.ua/portal/natural/Oeiet/2009_1/17pdf/9.pdf, accessed on 23 October 2009.

<http://www.thearling.com/text/whexcerpt/whexcerpt.htm>, accessed on 23 October 2009.

<http://www.twocrows.com/crm-dm.pdf>, accessed on 8 November 2009