

# Inculcation of knowledge discovery towards effective Customer Relationship Management (CRM)

\*Satyendra Kr. Singh

\*\*Sanchita Ghatak

---

## ABSTRACT

Knowledge discovery has been used to draw inferences from the data which can be used for formulation of strategies in the organisation. A number of companies have involved this concept to have effective customer relationship with their potential and existing customers.

Researches have proved that properly framed CRM has increased the profitability of the organisations. It is therefore crucial to discover the behaviour of the customers through knowledge discovery mechanism.

This paper discusses the concept of knowledge discovery and data mining and the importance of CRM as mined from the knowledge base and how the organisations can get benefited from it. It further discusses two cases where knowledge discovery is used to obtain valuable information about the customers.

## Introduction

Knowledge discovery (KD) is a concept of the field of computer science that describes the process of automatically searching large volumes of data for patterns that can be considered knowledge about the data. It is often described as deriving knowledge from the input data. This complex topic can be categorized according to 1) what kind of data is searched; and 2) in what form is the result of the search represented.

KD is also defined as "the non-trivial extraction of implicit, unknown, and potentially useful information from data". This sounds similar to data mining. However, there is a clear distinction between data mining and knowledge discovery. The knowledge discovery process takes the raw results from data mining (the process of extracting trends or patterns from data) and carefully and accurately transforms them into useful and understandable information. This information is not typically retrievable by standard techniques but is uncovered through the use of artificial intelligence (AI) techniques.

The most well-known branch of data mining is knowledge discovery, also known as Knowledge Discovery in Databases (KDD). Just as many other forms of knowledge discovery it creates abstractions of the input data. The knowledge obtained through the process may become additional data that can be used for further usage and discovery.

The amount of data being collected in databases today far exceeds our ability to reduce and analyze data without the use of automated analysis techniques. Many scientific and transactional business databases grow at a phenomenal rate. For instance, a single system, the astronomical survey application SCICAT, is expected to exceed three terabytes of data at completion. Knowledge discovery in databases (KDD) is the field that is evolving to provide automated analysis solutions.

KDD is a growing field and there are many knowledge discovery methodologies in use and under development. Some of these techniques are generic, while others are domain-specific. Large amounts of data

---

\* Lecturer, IISE, Lucknow mail2satyen@gmail.com

\*\* Sr. Lecturer, Jaipuria Institute of Management, Lucknow

are required to provide sufficient information to derive additional knowledge. Since large amounts of data are required, processing efficiency is essential. Accuracy is required to assure that discovered knowledge is valid. The results should be presented in a manner that is understandable by humans. One of the major premises of KDD is that the knowledge is discovered using intelligent learning techniques that sift through the data in an automated process. For this technique to be considered useful in terms of knowledge discovery, the discovered knowledge must be interesting; that is, it must have potential value to the user.

KDD provides the capability to discover new and meaningful information by using existing data. KDD quickly exceeds the human capacity to analyze large data sets. The amount of data that requires processing and analysis in a large database exceeds human capabilities, and the difficulty of accurately transforming raw data into knowledge surpasses the limits of traditional databases. Therefore, the full utilization of stored data depends on the use of knowledge discovery techniques.

The usefulness of future applications of KDD is far-reaching. KDD may be used as a means of information retrieval, in the same manner that intelligent agents perform information retrieval on the web. New patterns or trends in data may be discovered using these techniques. KDD may also be used as a basis for the intelligent interfaces of tomorrow, by adding a knowledge discovery component to a database engine or by integrating KDD with spreadsheets and visualizations.

Data mining is the process of extracting patterns from data. As more data is gathered, with the amount of data doubling every three years, data mining is becoming an increasingly important tool to transform these data into information. It is commonly used in a wide range of profiling practices, such as marketing, surveillance, fraud detection and scientific discovery.

While data mining can be used to uncover patterns in data samples, it is important to be aware that the use of non-representative samples of data may produce results that are not indicative of the domain. Similarly, data mining will not find patterns that may be present in the domain, if those patterns are not present in the sample being "mined". There is a tendency for insufficiently knowledgeable "consumers" of the results to attribute "magical abilities" to data mining, treating the technique as a sort of all-seeing crystal ball. Like any other tool, it only functions in conjunction with the appropriate raw material: in this case, indicative and representative data that the user must first collect. Further, the discovery of a particular pattern in a particular set of data does not necessarily mean that pattern is representative of the whole population from which that data was drawn. Hence, an important part of the process is the verification and validation of patterns on other samples of data.

The term data mining has also been used in a related but negative sense, to mean the deliberate searching for apparent but not necessarily representative patterns in large numbers of data. To avoid confusion with the other sense, the terms data dredging and data snooping are often used. Note, however, that dredging and snooping can be (and sometimes are) used as exploratory tools when developing and clarifying hypotheses.

### **Data mining process**

Data Mining is an analytic process designed to explore data (usually large amounts of data - typically business or market related) in search of consistent patterns and/or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data. The ultimate goal of data mining is prediction - and predictive data mining is the most common type of data mining and one that has

the most direct business applications. The process of data mining consists of three stages: (1) the initial exploration, (2) model building or pattern identification with validation/verification, and (3) deployment (i.e., the application of the model to new data in order to generate predictions).

### **Customer Relationship Management**

Customers pay money for the services or products they purchase. Hence, the companies need to develop effective plans to deal with this segment. Customer Relationship Management systems are also important to the top management because it provides crucial data like customer satisfaction and efficiency of service by the frontline crews. A piece of customer relationship management software will also be able to generate the needed reports for product development or new concepts. Furthermore, this system will also be a great help for the top management in deciding the company's future course of action, whether it involves phasing out one of the products on the shelves or making adjustments to one of the products sold.

This information is important in providing the customer the answer that he or she needs to resolve the issue without having to wait for a long time and without going to several departments. With just a few mouse clicks, a customer support representative for example can track the location of the customer's package or order. This is infinitely better than the cumbersome process of tracking shipments previously. Furthermore, the customer service representative will also be able to see the previous concerns of the customer. This is a great help especially if the customer is calling about the same issue since he or she will not have to repeat the story all over. This results in less time in resolving the issue, thus, higher productivity of the support staff.

The reports generated by CRM systems are also invaluable to the advertising and marketing planners, as they will be able to pinpoint which ideas works and which do not. Because of CRM systems, it is possible to release advertisements or plan marketing campaigns more in tune with the target market. This leads to more responses to the advertisement and a more effective marketing campaign.

### **Data Mining (Challenges & Opportunities) in CRM**

Although data mining is a very useful concept yet it has to face lot many challenges as discussed in the following section.

Non-trivial results almost always need a combination of DM techniques. Chaining/composition of DM, and more generally data analysis, operations is important. In order to analyze CRM data, one needs to explore the data from different angles and look at its different aspects. This should require application of different types of DM techniques and their application to different "slices" of data in an interactive and iterative fashion. Hence, the need to use various DM operators and combine (chain) them into a single "exploration plan".

There is a strong requirement for data integration before data mining. In both cases, data comes from multiple sources. For example in CRM, data needed may come from different departments of an organization. Since many interesting patterns span multiple data sources, there is a need to integrate these data before an actual data mining exploration can start.

Diverse data types are often encountered, which requires the integrated mining of diverse and heterogeneous data. In CRM, while dealing with this issue is not critical, it is nonetheless important.

Customer data comes in the form of structured records of different data types (e.g., demographic data), temporal data (e.g., weblogs), text (e.g., emails, consumer reviews, blogs and

chat-room data), (sometimes) audio (e.g., recorded phone conversations of service reps with customers), etc.

Highly and unavoidably noisy data must be dealt with. In CRM, weblog data has a lot of "noise" (due to crawlers, missed hits because of the caching problem, etc.). Other data pertaining to customer "touchpoints" has the usual cleaning problems seen in any business-related data.

Privacy and confidentiality considerations for data and analysis results are a major issue. In CRM, a lot of demographic data is highly confidential, as are emails and phone logs.

### **Data mining has the following applications in terms of CRM:**

Customer relationship management system to find out who are the best customers, which products are to be offered to which customers (direct marketing or customer acquisition), which customers are likely to end the relationship (customer churn), which customers are likely to not pay (also coined as fraud detection)

Decision support system applies to almost all areas, ranging from medicine over marketing to logistics. KDD applications aim at a data-driven justification of decisions by relating actions and outcomes.

Recommender systems rank objects according to user profiles. The objects can be, for instance, products as in the amazon internet shop, or documents as in learning search engines. KDD applications do not assume user profiles to be given but learns them from observations of user behavior.

Plant asset management system moves beyond job scheduling and quality control. The goal is to optimize the overall benefits of production.

In order to better understand the significance of data mining and how it is related to CRM, let us go through the following two brief cases.

#### **Case I**

Let us take into consideration a situation where a company intends to sustain high-profit, high-value, and low-risk customers. Such a segment is just 10-20 percent of the total customers yet account for 50-80 percent of profit. Cross-selling and up-selling is always a choice of the marketing department. The demographic data of such customers can be used to steer the advertising activities.

#### **Case II**

One of the major focuses for any bank is to sustain both the high-profit and low-profit customers. Further, it is one of the main activities of the bank to convert the low-profit customers into high-profit group. This requires analysis of the demography and other data that stimulate the customer behaviour. This is done by proper data mining of customer and capitalising through effective CRM.

#### **Conclusion**

Data is considered one of the important resources for any organisation. The analysis of this data reduces the time and helps in accurate prediction of the required information.

Knowledge discovery through data mining helps the organisations plan and execute CRM activities that can result in customer-retention and in turn generating higher profits in the long run.

## References

- *Balaji Padmanabhan, Alexander Tuzhilin (2002), Knowledge refinement based on the discovery of unexpected patterns in data mining*
- *Decision Support Systems, Volume 33, Issue 3, Pages 309-321*
- *E.W.T. Ngai, Li Xiu, D.C.K. Chau (2009), Application of data mining techniques in customer relationship management: A literature review and classification*
- *Expert Systems with Applications, Volume 36, Issue 2, Part 2, Pages 2592-2602*
- *Sung Ho Ha (2007), Applying knowledge engineering techniques to customer analysis in the service industry, Advanced Engineering Informatics, Volume 21, Issue 3, July 2007, Pages 293-301*
- *Chris Rygielski, Jun-Cheng Wang, David C. Yen (2002), Data mining techniques for customer relationship management, Technology in Society, Volume 24, Issue 4, Pages 483-502*
- *Sheng-Tun Li, Li-Yen Shue, Shu-Fen Lee (2006), Enabling customer relationship management in ISP services through mining usage patterns*
- *Expert Systems with Applications, Volume 30, Issue 4, Pages 621-632*
- *Emine Tu?, Merve ?akiro?lu, Ahmet Arslan (2006), Automatic discovery of the sequential accesses from web log data files via a genetic algorithm, Knowledge-Based Systems, Volume 19, Issue 3, Pages 180-186*
- *Iraj Mahdavi, Namjae Cho, Babak Shirazi, Navid Sahebjamnia (2008), Designing evolving user profile in e-CRM with dynamic clustering of Web documents*
- *Data & Knowledge Engineering, Volume 65, Issue 2, Pages 355-372*
- *Kumar S. & Layek S (2009), Data Mining and role of data mining in business*
- *Yichen Lin, Hwan-Yann Su, Shihen Chien (2006), A knowledge-enabled procedure for customer relationship management, Industrial Marketing Management, Volume 35, Issue 4, Pages 446-456*
- *YongSeog Kim (2006), Toward a successful CRM: variable selection, sampling, and ensemble, Decision Support Systems, Volume 41, Issue 2, Pages 542-553*
- *Zengyou He, Xiaofei Xu, Joshua Zhexue Huang, Shengchun Deng (2004), Mining class outliers: concepts, algorithms and applications in CRM*
- *Expert Systems with Applications, Volume 27, Issue 4, Pages 681-697*