International Journal of Technology Management & Humanities (IJTMH)
e-ISSN: 2454 – 566X, Volume 1, Issue 1, Ver. 1 (JUNE 2015), Pg 00-00 www.ijtmh.com

Application of Data Mining in Customer Relationships Management (CRM)

Author

Puneet Shukla

(Asst. Professor/Department of Computer Science/SR Group of Institutions, Lucknow)

Abstract : *Data mining is a continuous, iterative process that is the very core of business intelligence. It involves the use of data mining software, sound methodology and human creativity to achieve new insight through the exploration of data to uncover patterns, relationships, anomalies and dependencies. Data mining helps marketing professionals improve their understanding of customer behavior. In turn, this better understanding allows them to target marketing campaigns more accurately and to align campaigns more closely with the needs, wants and attitudes of customers and prospects. Data mining is part of a much larger series of steps that takes place between a company and its customers. The way in which data mining impacts a business depends on the business process, not the data mining process. Data mining extracts information from a database that the user did not know existed. Relationships between variables and customer behaviors that are non-intuitive are the jewels that data mining hopes to find.*

Key Words : *Data Mining, DBMS, CRM, Cycle Time, Business Process, Business Intelligence*.

1. Introduction:

The way in which companies interact with their customers has changed dramatically over the past few years. A customer's continuing business is no longer guaranteed. As a result, companies have found that they need to understand their customers better, and to quickly respond to their wants and needs. In addition, the time frame in which these responses need to be made has been shrinking. It is no longer possible to wait until the signs of customer dissatisfaction are obvious before action must be taken. To succeed, companies must be proactive and anticipate what a customer desires. In this situation key question is arises how to catch new customers.
There are some factors are working together to increase the complexity of customer relationships:

- Day by day arising marketing costs.
- New product offering.
- Compressed marketing cycle times.
- Competitors.

Successful companies need to react to each and every one of these demands in a timely fashion. The market will not wait for your response, and customers that you have today could vanish tomorrow. Interacting with your customers is also not as simple as it has been in the past. Customers and prospective customers want to interact on their terms, meaning that you need to look at multiple criteria when evaluating how to proceed. You will need to automate:

- The Right Offer
- To the Right Person
- At the Right Time
- Through the Right Channel

2. The Purpose of Data Mining.

Data mining helps marketing professionals improve their understanding of customer behavior. In turn, this better understanding allows them to target marketing campaigns more accurately and to align campaigns more closely with the needs, wants and attitudes of customers and prospects.
If the necessary information exists in a database, the data mining process can model virtually any customer activity. The key is to find patterns relevant to current business problems. Typical questions that data mining addresses include like:

- Which customers are most likely to drop their cellular phone service?
- What is the probability that a customer will purchase at least $100 worth of merchandise from a particular mail-order catalog?
- Which prospects are most likely to respond to a particular offer?

3. Data Mining, What is it?

Data mining (the analysis step of the "Knowledge Discovery in Databases" process, or KDD), an interdisciplinary subfield of computer science, is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use.
Simply put, data mining is a continuous, iterative process that is the very core of business intelligence. It involves the use of data mining software, sound methodology and human creativity to achieve new insight through the exploration of data to uncover patterns, relationships, anomalies and dependencies. We have achieved our reputation as the data mining industry's leading innovator by developing powerful, user friendly and affordable data mining technology, and by delivering comprehensive knowledge transfer to customers to enable them to take advantage of the business benefits data mining technology makes possible. For almost a decade we has taken the leadership role in broadening user understanding and acceptance of this technology as a highly value decision support system for a wide range of business applications in many different industries.

4. Task in Data mining

Data mining involves six common classes of tasks:

4.1. Variance Detection (Outlier/change/deviation detection) **–** The identification of unusual data records, that might be interesting or data errors that require further investigation.

4.2. Association rule learning (Dependency modelling) **–** 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.

4.3. Clustering **–** is the task of discovering groups and structures in the data that are in some way or another "similar", without using known structures in the data.

4.4. Classification **–** is the task of generalizing known structure to apply to new data. For example, an e-mail program might attempt to classify an e-mail as "legitimate" or as "spam".

4.5. Regression **–** attempts to find a function which models the data with the least error.

4.6. Summarization **–** providing a more compact representation of the data set, including visualization and report generation.

5. Application of Data Mining in Science and engineering

In recent years, data mining has been used widely in the areas of science and engineering, such as bioinformatics, genetics, medicine, education and electrical power engineering etc.

- Human rights

- Medical data mining
- Spatial data mining
- Temporal data mining
- Sensor data mining
- Visual data mining
- Music data mining
- Surveillance
- Pattern mining
- Subject-based data mining
- Knowledge grid

6. Free Open Source Data Mining Software and Applications

- ***Carrot2:*** Text and search results clustering framework.
- ***Chemicalize.org:*** A chemical structure miner and web search engine.
- ***ELKI:*** A university research project with advanced cluster analysis and outlier detection methods written in the Java language.
- ***GATE:*** a natural language processing and language engineering tool.
- ***KNIME:*** The Konstanz Information Miner, a user friendly and comprehensive data analytics framework.
- ***ML-Flex:*** A software package that enables users to integrate with third-party machine-learning packages written in any programming language, execute classification analyses in parallel across multiple computing nodes, and produce HTML reports of classification results.
- ***MLPACK library:*** a collection of ready-to-use machine learning algorithms written in the C++ language.
- ***Massive Online Analysis (MOA):*** a real-time big data stream mining with concept drift tool in the Java programming language.
- ***NLTK (Natural Language Toolkit*)**: A suite of libraries and programs for symbolic and statistical natural language processing (NLP) for the Python language.
- ***OpenNN:*** Open neural networks library.
- Orange: A component-based data mining and machine learning software suite written in the Python language.
- ***R:*** A programming language and software environment for statistical computing, data mining, and graphics. It is part of the GNU Project.
- ***SCaViS:*** Java cross-platform data analysis framework developed at Argonne National Laboratory.
- ***SenticNet API:*** A semantic and affective resource for opinion mining and sentiment analysis.
- ***Tanagra:*** A visualisation-oriented data mining software, also for teaching.
- Torch: An open source deep learning library for the Lua programming language and scientific computing framework with wide support for machine learning algorithms.
- ***UIMA:*** The UIMA (Unstructured Information Management Architecture) is a component framework for analyzing unstructured content such as text, audio and video – originally developed by IBM.
- ***Weka:*** A suite of machine learning software applications written in the Java programming language.

7. Commercial data-mining software and applications

- ***Angoss KnowledgeSTUDIO:*** data mining tool provided by Angoss.
- ***Clarabridge:*** enterprise class text analytics solution.
- ***HP Vertica Analytics Platform:*** data mining software provided by HP.
- ***IBM SPSS Modeler:*** data mining software provided by IBM.
- ***KXEN Modeler:*** data mining tool provided by KXEN.
- ***Grapheme:*** data mining and visualization software provided by iChrome.
- ***LIONsolver:*** an integrated software application for data mining, business intelligence, and modeling that implements the Learning and Intelligent OptimizatioN (LION) approach.
- ***Microsoft Analysis Services:*** data mining software provided by Microsoft.
- ***NetOwl:*** suite of multilingual text and entity analytics products that enable data mining.
- ***Oracle Data Mining:*** data mining software by Oracle.
- ***RapidMiner:*** An environment for machine learning and data mining experiments.

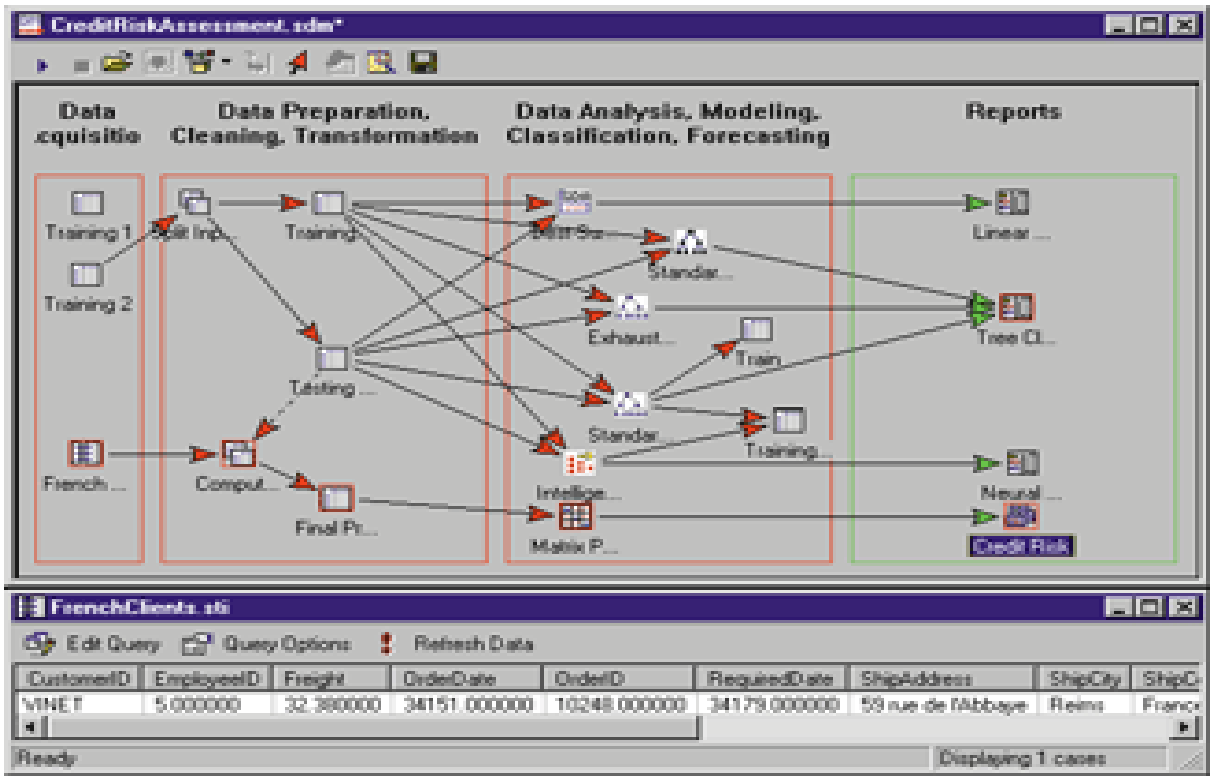
- ***SAS Enterprise Miner:*** data mining software provided by the SAS Institute.
- ***STATISTICA Data Miner:*** data mining software provided by StatSoft.
- ***Qlucore Omics Explorer:*** data mining software provided by Qlucore.

8. Process of Data Mining

The process of data mining consists of three stages:

(1) 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).

8.1. **Exploration.** This stage usually starts with data preparation which may involve cleaning data, data transformations, selecting subsets of records and - in case of data sets with large numbers of variables ("fields") - performing some preliminary feature selection operations to bring the number of variables to a



manageable range (depending on the statistical methods which are being considered). Then, depending on the nature of the analytic problem, this first stage of the process of data mining may involve anywhere between a simple choice of straightforward predictors for a regression model, to elaborate exploratory analyses using a wide variety of graphical and statistical methods in order to identify the most relevant variables and determine the complexity and/or the general nature of models that can be taken into account in the next stage.

8.2. **Model building and validation.** This stage involves considering various models and choosing the best one based on their predictive performance (i.e., explaining the variability in question and producing stable results across samples). This may sound like a simple operation, but in fact, it sometimes involves a very elaborate process. There are a variety of techniques developed to achieve that goal - many of which are based on so-called "competitive evaluation of models," that is, applying different models to the same data set and then comparing their performance to choose the best. These techniques - which are often considered the core of predictive data mining - include: Bagging (Voting, Averaging), Boosting, Stacking (Stacked Generalizations), and Meta-Learning.

8.3. **Deployment.** That final stage involves using the model selected as best in the previous stage and applying it to new data in order to generate predictions or estimates of the expected outcome.

9. How does data mining work?

While large-scale information technology has been evolving separate transaction and analytical systems, data mining provides the link between the two. Data mining software analyzes relationships and patterns in stored transaction data based on open-ended user queries. Several types of analytical software are available: statistical, machine learning, and neural networks. Generally, any of four types of relationships are sought:

9.1. **Classes**: Stored data is used to locate data in predetermined groups. For example, a restaurant chain could mine customer purchase data to determine when customers visit and what they typically order. This information could be used to increase traffic by having daily specials.
9.2. **Clusters**: Data items are grouped according to logical relationships or consumer preferences. For example, data can be mined to identify market segments or consumer affinities.
9.3. **Associations**: Data can be mined to identify associations. The beer-diaper example is an example of associative mining.
9.4. **Sequential patterns**: Data is mined to anticipate behavior patterns and trends. For example, an outdoor equipment retailer could predict the likelihood of a backpack being purchased based on a consumer's purchase of sleeping bags and hiking shoes.

Data mining consists of five major elements:

- Extract, transform, and load transaction data onto the data warehouse system.
- Store and manage the data in a multidimensional database system.
- Provide data access to business analysts and information technology professionals.
- Analyze the data by application software.
- Present the data in a useful format, such as a graph or table.

10. Business Process in Data Mining

Data mining is part of a much larger series of steps that takes place between a company and its customers. The way in which data mining impacts a business depends on the business process, not the data mining process. Take product marketing as an example. A marketing manager's job is to understand their market. With this understanding comes the ability to interact with customers in this market, using a number of channels. This involves a number of areas, including direct marketing, print advertising, telemarketing, and radio/television advertising, among others. Data mining, on the other hand, extracts information from a database that the user did not know existed. Relationships between variables and customer behaviors that are non-intuitive are the jewels that data mining hopes to find. And because the user does not know beforehand what the data mining process has discovered, it is a much bigger leap to take the output of the system and translate it into a solution to a business problem.

11. Data Mining and Customer Relationship Management

Customer relationship management (CRM) is a process that manages the interactions between a company and its customers. The primary users of CRM software applications are database marketers who are looking to automate the process of interacting with customers.

To be successful, database marketers must first identify market segments containing customers or prospects with high-profit potential. They then build and execute campaigns that favorably impact the behavior of these individuals. The first task, identifying market segments, requires significant data about prospective customers and their buying behaviors.

In theory, the more data the better. In practice, however, massive data stores often impede marketers, who struggle to sift through the minutiae to find the nuggets of valuable information. 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. This separation of the data mining and campaign management software introduces considerable inefficiency and opens the door for human errors. Tightly integrating the two disciplines presents an opportunity for companies to gain competitive advantage.

12. How Data Mining Helps Database Marketing

Data mining helps marketing users to target marketing campaigns more accurately; and also to align campaigns more closely with the needs, wants, and attitudes of customers and prospects. If the necessary information exists in a database, the data mining process can model virtually any customer activity. The key is to find patterns relevant to current business problems. Typical questions that data mining addresses include the following:

- which customers are most likely to drop their cell phone service?
- What is the probability that a customer will purchase at least $100 worth of merchandise from a particular mail-order catalog?
- which prospects are most likely to respond to a particular offer? Answers to these questions can help retain customers and increase campaign response rates, which, in turn, increase buying, cross selling, etc.

13. Business Intelligence Benefits

Data mining technology delivers two key business intelligence benefits:

1. Descriptive Function
2. Predictive Function

13.1. It enables enterprises, regardless of industry or size, in the context of defined business objectives, to **automatically** explore, visualize and understand their data, and to identify patterns, relationships and dependencies that impact on business outcomes (such as revenue growth, profit improvement, cost containment, and risk management) - a **descriptive** function.

13.2. It enables relationships uncovered and identified through the data mining process to be expressed as business rules, or predictive models. These outputs can be communicated in traditional reporting formats (presentations, briefs, electronic information sharing) to guide business planning and strategy. Also these outputs, expressed as programming code, can be deployed or "hard wired" into business operating systems to generate predictions of future outcomes, based on newly generated data, with higher accuracy and certainty - a **predictive** function.

For example, in the "CRM" arena a business can evaluate and develop a set of business intelligence rules about all aspects of its customer interactions. A simple example is modeling the likelihood of response to a specific solicitation of a new product or service. Based on these business rules, the business can target its marketing campaigns for maximum response to generate a desired level of response, revenue or profitability. Other typical "CRM" business examples would include:

- modelling customer acquisition (for targeted marketing and other CRM initiatives)
- assessing customer defection (for customer service and reclamation purposes)
- monitoring risk of loss (for customer scoring and credit approval decision making)

However, the reach of data mining technology extends far beyond "CRM" to encompass any process involving the acquisition, interpretation and acting on of data (internally or externally sourced). In the business domain this would include areas as diverse as internal audit and expense control through to research and development for new products or services. Using our data mining components, a wide range of solutions can be developed directly and easily as applications integrated with our data mining engine or as integrated components of mainstream "CRM" "sales force automation", "call center" and other enterprise applications - by third party solution providers or in the case of a large organization by their internal IT personnel working in collaboration and its partners.

14. Technology Issues

Data mining solutions are based on the implementation, through programming, of interfaces to generally available and privately developed algorithms, which enable the efficient exploration and organization of data. These algorithms support the identification of patterns, relationships and anomalies of potential interest to business decision-makers.

In addition to implementing these algorithms in a user accessible method, data mining technology also requires an understanding of various databases and implementation of data mining solutions to take advantage of features of such databases (if any) that make data mining tasks more efficient over large volumes of data.

In addition to algorithm implementation issues, key considerations relative to data mining solutions are data preparation issues and ensuring scalability and performance over very large volumes of data.

15. Conclusion

Data mining is primarily used today by companies with a strong consumer focus - retail, financial, communication, and marketing organizations. It enables these companies to determine relationships among "internal" factors such as price, product positioning, or staff skills, and "external" factors such as economic indicators, competition, and customer demographics. And, it enables them to determine the impact on sales, customer satisfaction, and corporate profits. Finally, it enables them to "drill down" into summary information to view detail transactional data.

With data mining, a retailer could use point-of-sale records of customer purchases to send targeted promotions based on an individual's purchase history. By mining demographic data from comment or warranty cards, the retailer could develop products and promotions to appeal to specific customer segments.

Comprehensive data warehouses that integrate operational data with customer, supplier, and market information have resulted in an explosion of information. Competition requires timely and sophisticated analysis on an integrated view of the data. However, there is a growing gap between more powerful storage and retrieval systems and the users' ability to effectively analyze and act on the information they contain. Both relational and OLAP technologies have tremendous capabilities for navigating massive data warehouses, but brute force navigation of data is not enough. A new technological leap is needed to structure and prioritize information for specific end-user problems. The data mining tools can make this leap. Quantifiable business benefits have been proven through the integration of data mining with current information systems, and new products are on the horizon that will bring this integration to an even wider audience of users.

Acknowledgement

References

[1] Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic (1996). "From Data Mining to Knowledge Discovery in Databases (PDF). Retrieved 17 Feb 2015.

[2] Data Mining Curriculum ACM SIGKDD. 2006-04-30. Retrieved 2015-01-27.

[3] Clifton, Christopher (2010). Encyclopedia Britannica: Definition of Data Mining. Retrieved 2015-05-09.

[4] Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome (2009). "The Elements of Statistical Learning: Data Mining, Inference, and Prediction

[5] Han, Jiawei; Kamber, Micheline (2001). *Data mining: concepts and techniques*. Morgan Kaufmann. p. 5. ISBN 9781558604896.

[6] OKAIRP 2005 Fall Conference, Arizona State University

[7] Berry, M., J., A., & Linoff, G., S., (2000). Mastering data mining. New York: Wiley.

[8] Edelstein, H., A. (1999). Introduction to data mining and knowledge discovery (3rd ed). Potomac, MD: Two Crows Corp.

[9] Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (1996). Advances in knowledge discovery & data mining. Cambridge, MA: MIT Press.

[10] Han, J., Kamber, M. (2000). Data mining: Concepts and Techniques. New York: Morgan-Kaufman.

[11] Hastie, T., Tibshirani, R., & Friedman, J. H. (2001). The elements of statistical learning : Data mining, inference, and prediction. New York: Springer.

[12] Pregibon, D. (1997). Data Mining. Statistical Computing and Graphics, 7, 8.

[13] Weiss, S. M., & Indurkhya, N. (1997). Predictive data mining: A practical guide. New York: Morgan-Kaufman.

Author Profile

Author (Puneet Shukla) has completed B.Tech (Computer Science) in 2010 and perusing M.Tech. He has more than 4 years of experience in academics as he is currently working as Asst professor in SR Group of Institutions, Lucknow, department of Information Technology. With the academics author has publish 2 International Research Paper, 3 national Research Papers, attended 2 National Conferences and Seminars, 1 International Seminars. Apart from SR Group of Institutions, Lucknow, he is Visiting Faculty of Study Centers of Sikkim Manipal University. He has taught subjects as Parallel Algorithm, DBMS, Object Oriented System, Web Technology, and Operating System.