

CRM WITH DATA MINING & WAREHOUSE: "OPTIMIZES CUSTOMER INSIGHT"

Dr. Vikas Saraf MBA, Ph.D, FDP – IIM, Indore *

Pawan Thakur**

Ms. Lata Yadav***

ABSTRACTS

Understanding customer behavior is important to adjusting business strategies, increasing revenues, identifying new opportunities and expanding business activities. The vital importance of such knowledge of these objectives is not new; in fact it's always been a fixture of business success, what's new is that many organizations have an impressive variety of data & information resources that leads to reveal about customer behavior, purchasing decision, price sensitivity, attitudes, etc. Diverse businesses and leading enterprises, now-a-days uses data warehouses with customer focus that sets them apart from their traditional operational databases. The data mining and data warehouses have become the bases on which specific business aims can be achieved & predictive model can be built. Generally, data mining and warehousing sometimes called data or knowledge discovery are the process of analyzing data from different perspectives and summarizing it into useful information where information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles,

* (Former Joint Director-ICWAI), Professor & Director, Vidyasagar Institute of Management (VIM), Near Awadhपुरi, BHEL, Bhopal (M.P) – 462 022.(India)

** Assistant Professor, Department of Computer Application & IT , Vidyasagar Institute of Management (VIM), Near Awadhपुरi, BHEL, Bhopal (M.P)- 462 022 (India)

*** Assistant Professor, Vidyasagar Institute of Management (VIM), Near Awadhपुरi, BHEL, Bhopal (M.P)- 462 022 (India).

categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases. This work presents the Data Mining and Warehousing in CRM, integration of CRM with Data Mining & Warehouse by creating a customer database, explain the whole Data Mining process for CRM and facilitate of Data Mining & Warehousing for organization.

Keywords: *Mining, *Warehouse, *CRM, *SPSS, *Operational, *Analytical, *Collaborative.

1.1 Introduction

"Data Mining is the process of extracting & presenting new knowledge, previously undetectable, selected from databases for actionable decision [1, 3]." Data Mining is the process of extracting valid, previously unknown & ultimately comprehensible information from a large database and using it to solve business problems and to make crucial business decisions. Data mining & knowledge discovery are receiving increasing attention in the business & technological press, among industry analysts, & among corporate management.

The data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process [11]. A data warehouse can be used to analyze a particular subject area. For example, "sales" can be a particular subject. A data warehouse integrates data from multiple data sources. For example, source A and source B may have different ways of identifying a product, but in a data warehouse, there will be only a single way of identifying a product [4]. Historical data is kept in a data warehouse. For example, one can retrieve data from 3 months, 6 months, 12 months, or even older data from a data warehouse. This contrasts with a transactions system, where often only the most recent data is kept. For example, a transaction system may hold the most recent address of a customer, where a data warehouse can hold all addresses associated with a customer.

CRM is the abbreviation for Customer Relationship Management. *CRM entails all aspects of interaction that a company has with its customer, whether it is sales or service-related.* CRM is often thought of as a business strategy that enables businesses to: Understand the customer, Retain customers through better customer experience, Attract new customer, Win new clients and contracts, and Increase profitably, Decrease customer management costs [1, 3, 7].

In the first stage using data mining, in Data Collection, individual sites collected data used to make simple calculations such as summations or averages. Information generated at this step answered business questions related to figures derived from data collection sites, such as total revenue or average total revenue over a period of time. Specific application programs were created for collecting data and calculations.

The second step using data warehousing, in Data Access, used databases are stored data in a structured format. At this stage, company-wide policies for data collection and reporting of management information were established. Because every business unit conformed to specific requirements or formats, businesses could query the information system regarding branch sales during any specified time period.

1.2 Data Mining & Warehousing in CRM

In CRM, data mining is frequently used to assign a score to a particular customer or prospect indicating the likelihood that the individual behaves the way we want. For example: a score could measure the propensity to respond to a particular offer or to a competitor's product. It is also frequently used to identify a set of characteristics (called a profile) that segments customers into groups with the similar behavior such as buying a particular product [9].

The first and simplest analytical step in data mining is to describe the data. For example: organization can summarize data's statistical attributes (such as Means & Standard Deviations), visually review data using charts and graphs. But data description alone cannot provide an action plan. Companies must build a predictive model based on patterns determined from known results & then test those model results outside the original sample. A good model should never be confused with reality, but it can be a useful guide to understanding business. Data mining can be used for both classifications & regression problems. In classification problems we are predicting what category something falls into - For example:- whether or not a person is a good credit risk or which of several offers someone is most likely to accept. In regression problems, we are predicting a number, such as the probability that a person will respond to an offer. Today detailed customer interaction data is abundant. We might have data about browsing behavior, returns, complaints, wishes, gifts and more [4, 5, 6].

How many businesses are truly using this data effectively?

The reason for this paradox is that technology for generating, capturing and storing data has far outpaced the human capacity to understand, analyze and exploit it for maximum impact. Data

Mining Technology, which focuses on identifying interesting patterns and developing predictive models from data, has the greatest potential for enabling businesses to leverage data resources for strategic business success.

1.3 Integration of CRM with Data Mining & Warehouse: Creating a Customer Database

Customer relationship management is a process whose objective is to enhance customer loyalty. This process consists of the following:

- Creation and management of data mines & warehouses
- Development of appropriate organizational structures
- Investment in technology
- People development.

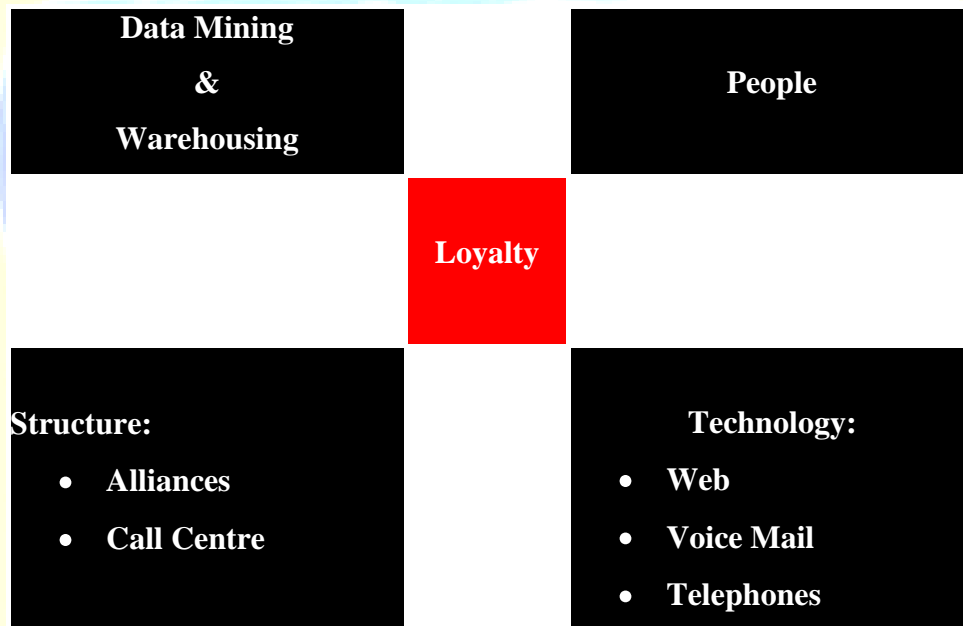


Fig. 1.1 Integration of CRM with Data Warehouse

CRM process depends on data. Single operation focused integrated logical database, data warehousing, data mining, decision support system (DSS), campaign management tools as well as call center software and hardware[6,7,4]. A necessary step to a complete CRM solution is the construction of a customer database or information file. This is the foundation for any customer relationship management activity. For web based businesses, constructing a database should be a relatively straightforward task, as the customer transaction & contact information is accumulated as a natural part of the interaction with customers. For existing companies that have not

previously collected much customer information, the task will involve seeking historical customer contact data from internal sources such as accounting & customer service.

What should be collected for the database? Ideally, the database should contain information about the following:

- (i). **Transactions.** This should include a complete purchase history with accompanying details (price paid, delivery date, location etc.).
- (ii). **Customer Contacts.** Today, there is an increasing number of customer contact points from multiple channels and contexts. This should not only include sales calls and service requests, but any customer-or-company-initiated contact.
- (iii). **Descriptive Information.** This is for segmentation and other data analysis purposes.
- (iv). **Response to Marketing Stimuli.** This part of the information file should contain whether or not the customer responded to a direct marketing initiative, a sales contact, or any other direct contact. The data should be represented over time [10].

Companies have traditionally used variety of methods to construct their databases. Durable goods manufactures utilize information from warranty cards for basic descriptive information. Unfortunately, response rates to warranty cards leave big gaps in the databases. Service businesses are normally in better shape since the nature of the product involves the kind of customer-company interaction that naturally leads to better data collection. For examples, banks have been in the forefront of CRM activities for a number of years. Telecom-related industries (long distance, wireless, cable services) similarly have a large amount of customer information.

There are following three types of interactive CRM technologies, as mentioned in fig. 1.2:

1. Operational CRM. Most of the CRM software focused on simplifying the organization and management of customer information. Such software, called '**Operational CRM**', focuses on creating a customer database that presents a consistent picture of the customer's relationship with the company and providing that information in specific applications. This include sales force automation and customer service applications, in which the company "touches" the customer [9].

Operational
CRMAnalytical
CRMCollaborative
CRM

CUSTOMER

Interaction of CRM Technologies

Fig. 1.2 Interaction of CRM Technologies

2. Analytical CRM. Data mining is a process that uses a variety of data analysis and modeling techniques to discover patterns and relationships in data that are used to understand what customers want and predict what they will do. Data mining can help to select the right prospects on whom to focus, offer the right additional products to existing customers and identify good customers who may be about to leave. These results are in improved revenue because of a greatly improved ability to respond to each individual contact in the best way and reduced costs due to properly allocated resources. CRM applications that use data mining are called '**Analytical CRM**'.

3. Collaborative CRM. Data mining also frequently used to identify a set of characteristics that segments customers into groups with similar behaviors, such as buying a particular product. A special type of classification can recommend items based on similar interests held by groups of customers. This is called '**Collaborative CRM**'.

Data mining can improve profitability in each stages of customer life cycle when integrated with operational, analytical or collaborative CRM systems or implement it as independent applications.

1.4 Data Mining Process for CRM

For successful operational CRM, organizations need to maintain a strict data flow framework within their organization. A broad framework for building data flow framework within any organization would consist of the following processes as shown in Fig. 1.3[9]:

- 1. Identification of the Objective.** To identify the core areas in which organization would need to implement effective customer centric views.
- 2. Data Preparation.** To identify the different sources of customer related data from within the organization or outside. The common sources of data for organizations are from their transactional systems, marketing data, sales record and market research data. Secondary data

sources from outside the organization could also be used. Three-sources of customer data are most critical to data mining efforts toward better understanding of behavior, which are as follows: -

(i). **Demographic Data.** Direct marketers have employed data about age, geographical location, & income for many years to target specific groups of customers. The goal has been to use data to aim promotional campaigns at groups with particular interests.

(ii). **Transaction Data.** This resource provides concrete data about what your customers are purchasing. Going beyond general demographics information, transaction data is essential in helping to predict future purchases & target promotional campaigns more effectively. In addition to the value of the transaction itself, this data also reveals key information about time, location, & other factors related to the transaction.

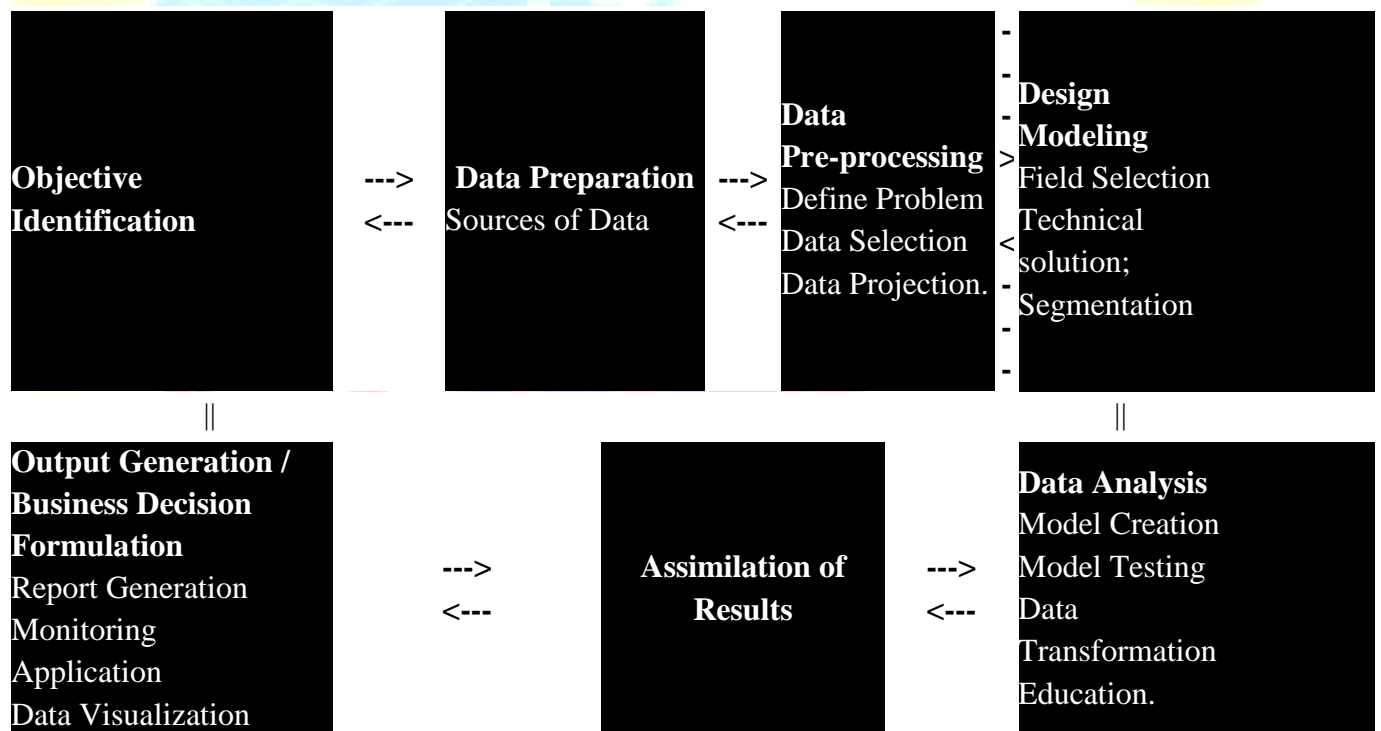


Fig. 1.3 Data Mining Process For CRM

(iii). **Online Interaction Data / Online Transaction Processing (OLTP).** The dominant form of data here is 'Internet Clickstream Data', although we must also include interactions that occur through wireless devices, cable television & more. This resource can provide deeper information than transaction data; it provides a window on customers' decision processes and the navigational steps they took to find what they desired [5]. Online interaction data records every page that the

customer saw leading up to a decision. Therefore, we know not only what they purchased (or didn't purchase) but we also have strong evidence of how they arrived at that decision.

3. Data Processing. For analysis, data collected from different sources needs to be collected & pooled together into a single repository. However, before any organization can do this, the data collected from different sources has to be made consistent, since the data could be collected from systems that are running on different platforms, architectures or application systems.

4. The Model Design. It entails selecting an appropriate data mining algorithm to be applied to the data. Then, segmentation to apply to the data minimum will usually entail breaking the data out into a training set and one or more test sets. Segmentation might involve using clustering techniques to break the data into separate subsets based on common characteristics and then analyzing each segment separately.

Note:- "Don't confuse- Segmentation with Clustering".

"**Segmentation** refers to the general problem of identifying groups that have common characteristics". "**Clustering** is a way to segment data into groups that are not previously defined, whereas classification is a way to segment data by assigning it to groups that are already defined".

"Statistical Package for the Social Sciences (SPSS) provides solutions that discover what customer want and predict what they will do. The company delivers solutions at the intersection of customer relationship management and business intelligence that enables its customers to interact with their customers more profitably. SPSS solutions integrate and analyze marketing, customer & operational data in key vertical markets worldwide including: - telecommunications, healthcare, banking, finance, insurance, manufacturing, retail, customer packaged goods, market research & the public sector".

5. Data Analysis. Data which has been collected may then be subjected to different data mining applications depending on these business requirements. Thus, phase involves a further preparatory activity (data transformation) to reorganize the data to best match the selected algorithm and the business problem. The selected data-mining tool is applied to the data, & this typically involves creating a model using the training set of data & then verifying the model with at least one separate set of test data. The model's accuracy & validity can then be evaluated.

It is very likely that the initial model will not meet the goals of the data mining exercise & that much iteration will be necessary, especially among the Model design & Data Analysis phases. This will involve trying out different data mining techniques or parameters on different subsets of the data before arriving at a successful outcome.

6. Assimilation of Results. Results obtained from the analysis could then be analyzed for effectiveness or applicability in the business process. Systems working without preconceived hypotheses, as in Data Mining Applications, could throw up results that might not be of direct relevance to organization process. All results have to be analyzed to identify those that may be converted into strategic knowledge for the business objective. This is the stage at which organizations would be able to formulate different value propositions for their customers.

7. Output Generation / Business Decision Formulation. Depending on the result obtained from the Analysis state, a business decision might be taken. This could be an automated business decision for real-time application system, or a long-term strategic decision for any organization depending on the requirement. This stage ideally represents the selection of a value proposition or a set of value propositions that the organization might want to present to their customers.

1.5 Facilitate of Data Mining & Warehousing for Organization

Data mining is a broad technology that can potentially benefit any functional areas within a business where there is a major need or opportunity for improved performance and where data can impact the performance improvement is available for analysis. The applications of the technology of data warehousing and / or data mining towards solving business problems like: - Target Marketing, Customer Retention, Fraud Detection, Customer Segmentation, Credit Risk assessment, Security Management, Resource Management, Customer Profitability Analysis, Customer Service Automation, Campaign Management, Product Planning, Cross-Selling / Up-Selling, Distribution Channel Management, Inventory Control, Relationship Marketing etc. Customer relationship management in its broadest sense simply means managing all customer information and interactions. In practice, this requires using information about customers and prospects to more effectively interact with customers in all stages of relationship with them. Data Warehousing / Data Mining facilitate the organization in all the 3-stages (shown in Fig.1.4) of customer life cycle i.e.;

- Acquiring new customers via data mining
- Increasing the value of existing customers: cross-selling via data mining

- Retaining good customers via data mining.

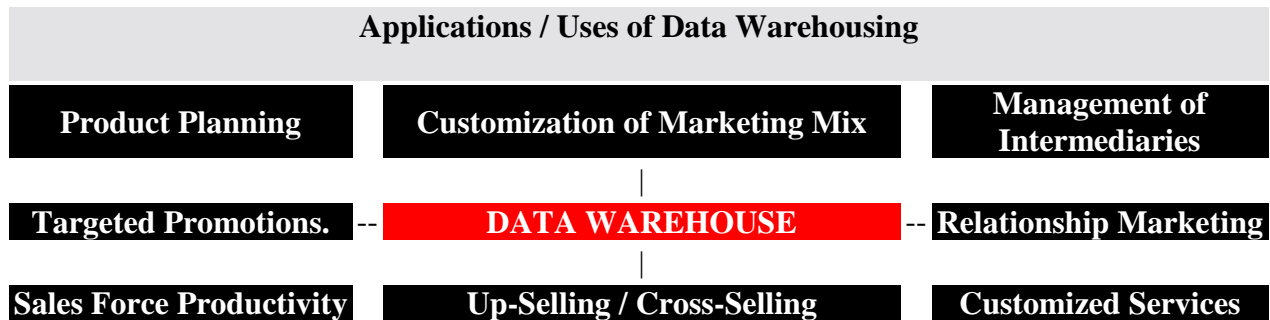


Fig. 1.4 Data Warehousing / Data Mining Facilitate the Organization

CRM helps companies improve profitability of their interactions with customers, while at the same time, makes the interactions appear friendlier through individualization. To succeed with CRM, companies need to match products and campaigns to prospects and customers – in other words, to intelligently manage the customer life cycle.

1.4 Conclusion

There is the growing significance of data mining & warehousing in customer satisfaction & business development. In order to be effective, data mining has to be more intelligent & offer information of the customer in real time. The data mining should help the organization to disseminate information on customers to everybody in the organization; which should facilitate each person's functioning & also make him/her customer responsive. Companies have used & are using data mining to support their sales & service staff in particular. They have also supported their sales & service staff with advance technology, which in turn has helped them to use the data for the purposes of developing a customer offer. Smart Cards & Credit Cards are good examples of the customer sharing his data with the organization.

-----XXX-----

References

1. Brown, Stanley A (n.d.), "*Customer Relationship Management*", John Wiley & Sons Ltd; Canada.
2. Fayyad, Usama (2003), "*Optimizing Customer Insight*", Intelligence Enterprise.
3. Edelstein Herb, President, "*Two crows corporation,- Building Profitable CRM with Data Mining*", <http://www.computerhistory.org>.
4. Kharbanda, Vikas & Dasgupta, Parthasarathi (2010), "*Data mining for Customer Relationship Management*", <http://www.yasni.com>.
5. Microsoft Press (n.d), "*Role of Database servers in CRM applications*".
6. Mohamed, H. Peeru & Sagadevan A (2002), "*Customer Relationship Management - A Step-by-Step Approach*", vikas publishing House Pvt. Ltd; New Delhi.
7. Saarevirta, Gary (n.d.), "*A characteristics of Data Mining Technologies & Processing*", Information Discovery Inc.
8. Saraf, Vikas (2003), "*Understanding & Implementing CRM: A Key Component for Tomorrows' Business Leader's*", Indian Journal of Marketing New Delhi Vol.-1,.
9. Saxena, Rajan, (2009), "*Marketing Management*", Tata McGraw-Hill Publication Ltd; New Delhi.
10. Teklitz, Frank & McCrathy, Robert L. (1993), "*Analytical Customer Relationship Management*", Siebel Corporation.
11. Winer, Russell S. (2001), "*A framework for Customer Relationship Management*", California Management Review Vol. 43, No. 4.

-----XXX-----