

Data Mining: A Competitive Tool in the Banking and Retail Industries

The banking industry around the world has undergone a tremendous change in the way business is conducted. Leading banks are using Data Mining (DM) tools for customer segmentation and profitability, credit scoring and approval, predicting payment default, marketing, detecting fraudulent transactions, etc. This article provides an overview of the concept of DM.

Data might be one of the most valuable assets of any corporation—but only if it knows how to reveal valuable knowledge hidden in raw data. Data mining allows to extract diamonds of knowledge from the historical data, and predict outcomes of future situations. It helps optimise business decisions, increase the value of each customer and communication, and improve customer satisfaction.

Data mining is the process of extracting previously unknown information, typically in the form of patterns and associations, from large databases. Today, organisations are realising the numerous advantages that come with data mining. It is a valuable tool—by identifying potentially useful information from the large amounts of data collected, an organisation can gain a clear advantage over its competitors.

Data mining can help companies in better understanding of the vast volume of data collected by the CRM systems. In the past few years, many organisations (especially retailers and banks) have recognised the vital importance of the information they have on their customers. Data mining, can identify products that are often purchased together, which can help build product bundles that are more likely to be successful.

Enterprises can also use data mining to minimise purchasing costs; score suppliers by rating the quality of their goods and services;

identify the most effective promotions; and address numerous other organisational needs, including fraud detection, failure analysis, predictive maintenance, risk management and demand forecasting. As is evident, data mining solutions have a wide variety of applications across industries.

Today, data mining is being used by several industries including banking and finance, retail, insurance, telecommunications, etc. Other possible applications for data mining include database marketing, sales forecasting, call behaviour analysis and churning management in telecommunications; forecasting of demand for utilities, such as energy and water; simulation of chemical and other process reactions; finding critical factors in discrete manufacturing (aerospace, automobile, electronics); CPU usage and forecasting. Data mining is often referred to as 'analytical intelligence'.

Several recent trends have increased the interest in data mining, including the declining cost of data storage and the increasing ease of collecting data, the development of robust and efficient machine-learning algorithms to process data, and the declining cost of computational power. With greater data storage capabilities and declining costs, data mining has offered organisations a new way of doing business. Data mining can help organisations better understand their business, be able to better serve their customers, and increase the effectiveness of the organisation in the long run. To stay competitive and keep abreast of industry dynamics, enterprises are using IT to generate, store and analyse mass-produced data not only for operational purposes



—Dr. Madan Lal Bhasin

(The author is Head, Accounting Department, Mazoon College, Muscat, Sultanate of Oman. He can be reached at madan.bhasin@rediffmail.com)

but also to enable strategic decision making.

The banking sector consists of public sector, private sector and foreign banks, apart from smaller regional and cooperative banks. In the market, various IT-based banking products, services and solutions are available. The most common of them are Phone Banking; ATM facility; Credit, Debit and Smart Cards; Internet Banking & Mobile Banking; SWIFT Network & INFINET Network; connectivity of bank branches to facilitate anywhere banking.

In addition to the above, software are also available to support various banking requirements. For instance, MIS helps banks to generate RBI reports like SLR, CRR (Cash Reserve Ratio) etc.; NPA Management, Asset & Liability Management, GAP and Trend Analysis; Branch-wise Profitability Assessment; Intelligent Balance Sheet Analyser; Relational Database Management Systems for management of data mining from data warehouse.

The banking sector is on the cusp of revolutionary change in the way it functions and delivers its services to customers. The traditional 'brick-and-mortar' bank branches are getting

data at the point-of-transaction itself, banks are amassing a huge volume of data that has immense business values buried in it and can provide valuable insights into the banking habits of customers of varying profiles. A substantial part of budget allocation, out of the total expenditures at banks and insurance companies, is being done for IT implementation. Customer Relationship Management (CRM) has become a rage in all service industries, and encompasses the data mining activities that a service provider undertakes to understand its customers.

What is Data Mining?

Data mining refers to computer-aided pattern discovery of previously unknown interrelationships and recurrences across seemingly unrelated attributes in order to predict actions, behaviours and outcomes. Data mining, in fact, helps to identify patterns and relationships in the data. Because data mining is a relatively new concept, it has been defined in various ways by various authors in the recent past. Box 1 provides a few definitions from literature on what data mining is, and how it can be beneficial to an organisation.

Box 1: Data Mining Defined Throughout Literature

1. Data mining is defined as the process of extracting previously unknown, valid, and actionable information from large databases and then using the information to make crucial business decisions – Cabena et al.
2. Data mining is described as the automated analysis of large amounts of data to find patterns and trends that may have otherwise gone undiscovered — Fabris.
3. The objective of data mining is to identify valid, novel, potentially useful, and understandable correlations and patterns in existing data — Chung and Grey

networked and becoming an integral part of an enterprise-wide banking platform, called 'Core Banking Solutions' (CBS). Implementation of CBS is enabling banks to introduce several innovative customer-friendly facilities and services, like anywhere banking, multi-city at par cheques, Internet banking, etc.

With banking transactions moving to an online platform that captures transactional

It may be noted at the outset that definitions of data mining given by these authors are somewhat different, but all have a common idea: to extract important information from existing data and enable better decision-making throughout an organisation. Not only can data mining improve decision-making by searching for relationships and patterns from the extensive data collected by organisations, it can

also reduce information overload. Data mining enables an organisation to focus on the most important information in the database, which allows managers to make more knowledgeable decisions by predicting future trends and behaviours. Before data mining was utilised, managers were not as capable of making such informed decisions because searching through large amounts of data was too expensive and time-consuming.

Data mining, as is evident from its key elements, typically involves the use of predictive modelling, forecasting and descriptive modelling techniques. By using these techniques, an organisation can proactively manage customer retention, identify cross-sell and up-sell opportunities, profile and segment customers, set optimal pricing policies, and objectively measure and rank which suppliers are best suited for their needs.

Data Mining Operations

Data mining operations are classified in different ways. A few categorisations of data mining operations are: clustering/segmentation, visualisation, predictive modelling, link analysis, deviation detection, dependency modelling, and summarisation. With most analysis tools, users know what questions they want to ask before using the tool. With data mining, users do not necessarily need to know what questions to ask; the software uncovers patterns for the users. Data visualisation tools allow users to 'see' patterns in the data. These tools allow users to build quite complex, three-dimensional representations of the data.

Many of the methodologies that are commonly described as forming part of a standard data mining tool kit have been developed either by the machine-learning community or by statistical data analysts. In fact, it is really difficult to specify any particular technique as being unique to the data mining community. Therefore, any listing of data analysis methodologies as data mining techniques would appear 'subjective'

and somewhat arbitrary. Some widely used techniques in data mining include artificial neural networks, genetic algorithms, K-nearest neighbour method, decision trees, and data reduction.

DM in the Banking Industry

The banking industry across the world has undergone tremendous changes in the way the business is conducted. With the recent implementation, greater acceptance and usage of 'electronic' banking, the capturing of transactional data has become easier and, simultaneously, the volume of such data has grown considerably. It is beyond human capability to analyse this huge amount of raw data and to effectively transform the data into useful knowledge for the organisation. The enormous amount of data that banks have been collecting over the years can greatly influence the success of data mining efforts. By using data mining to analyse patterns and trends, bank executives can predict, with increased accuracy, how customers will react to adjustments in interest rates, which customers will be likely to accept new product offers, which customers will be at a higher risk for defaulting on a loan, and how to make customer relationships more profitable.

The banking industry is widely recognising the importance of the information it has about its customers. Undoubtedly, it has among the richest and largest pool of customer information, covering customer demographics, transactional data, credit cards usage pattern, and so on. As banking is in the service industry, the task of maintaining a strong and effective CRM is a critical issue. To do this, banks need to invest their resources to better understand their existing and prospective customers. By using suitable data mining tools, banks can subsequently offer 'tailor-made' products and services to those customers.

There are numerous areas in which data mining can be used in the banking industry, which include customer segmentation and

profitability, credit scoring and approval, predicting payment default, marketing, detecting fraudulent transactions, cash management and forecasting operations, optimising stock portfolios, and ranking investments. In addition, banks may use data mining to identify their most profitable credit card customers or high-risk loan applicants. There is, therefore, a need to build an analytical capability to address the above-stated issues and data mining attempts to provide the answer. Following are some examples of how the banking industry has been effectively utilising data mining in these areas.

Marketing: One of the most widely used areas of data mining for the banking industry is marketing. The bank's marketing department can use data mining to analyse customer databases and develop statistically sound profiles of individual customer preferences for products and services. By offering only those products and services that customers really want, banks can save substantial money on promotions and offerings that would otherwise be unprofitable. Bank marketers, therefore, need to focus on their customers by learning more about them. Bank of America, for instance, uses database marketing to improve customer service and increase profits. By consolidating five years of customer history records, the bank was able to market and sell targeted services to customers.

'Cross-selling' is another marketing area where data mining can be extensively used. Here, a service provider makes it attractive for a customer to buy additional products or services with the same business. The more products and services a bank can provide for customers, the more likely the bank is to retain those customers. For example, several leading private and foreign banks in India (ICICI, HSBC, etc.) use data mining to find customers with demand deposit accounts who may be interested in a home equity loan. A model is built of the customers who already have home equity loans and this model is used to pinpoint other customers who may also be

interested. Another example is Bank of America, which has recently completed a project with IBM's data mining tools to search its database of corporate clients and try to figure out what products the clients may need next.

Risk Management: Data mining is widely used for risk management in the banking industry. Bank executives need to know whether the customers they are dealing with are reliable or not. Offering new customers credit cards, extending existing customers lines of credit, and approving loans can be risky decisions for

Data mining refers to computer-aided pattern discovery of previously unknown interrelationships and recurrences across seemingly unrelated attributes in order to predict actions, behaviours and outcomes. It is used by several industries, including banking & finance, retail, insurance, telecommunications, etc.

banks if they do not know anything about their customers. Data mining, however, can be used to reduce the risk of banks that issue credit cards by determining those customers who are likely to default on their accounts. An example was reported in the press of a bank discovering that cardholders who withdrew money at casinos had higher rates of delinquency and bankruptcy. It is a common practice on the part of banks to analyse customers' transaction behaviours in their deposit accounts to determine their probability of default in their loan accounts.

Credit scoring, in fact, was one of the earliest financial risk management tools developed. Credit scoring can be valuable to lenders in the banking industry when making lending decisions. Lenders would not have expanded the number of loans they give out without having an accurate, objective, and controllable risk assessment tool. The examples of both a 'good' and 'bad' loan applicant's histories can be used to develop a profile for a good and bad

'new' loan applicant.

Data mining can also derive the credit behaviour of individual borrowers with instalment, mortgage and credit card loans, using characteristics such as credit history, length of employment and length of residency. A score is thus produced that allows a lender to evaluate the customer and decide whether the person is a good candidate for a loan, or if there is a high risk of default. Customers who have been with the bank for longer periods of time, remained in good standing, and have higher salaries/wages, are more likely to receive a loan than a new customer who has no history with the bank, or who earns low salaries/wages. By knowing what the chances of default are for a customer, the bank is in a better position to reduce the risks.

Fraud Detection: Another popular area where data mining can be used in the banking industry is in fraud detection. Being able to detect fraudulent actions is an increasing concern for many businesses; and with the help of data mining more fraudulent actions are being detected and reported. Two different approaches have been developed by financial institutions to detect fraud patterns. In the first approach, a bank taps the data warehouse of a third party (potentially containing transaction information from many companies) and uses data mining programs to identify fraud patterns. The bank can then cross-reference those patterns with its own database for signs of internal trouble. In the second approach, fraud pattern identification is based strictly on the bank's own internal information. Most of the banks are using a 'hybrid' approach.

One system that has been successful in detecting fraud is Falcon's 'fraud assessment system'. It is used by nine of the top ten credit card issuing banks, where it examines the transactions of 80 per cent of cards held in the US. Mellon Bank also uses data mining for fraud detection and is able to better protect itself and its customers' funds from potential credit card fraud.

Customer Acquisition and Retention: Not only can data mining help the banking industry to gain new customers, it can also help retain existing customers. Customer acquisition and retention are very important concerns for any industry, especially the banking industry. Today, customers have so many opinions with regard to where they can choose to do their business. Executives in the banking industry, therefore, must be aware that if they are not giving each customer their full attention, the customer can simply find another bank that will. Data mining can also help in targeting 'new' customers for products and services and in discovering a customer's previous purchasing patterns so that the bank will be able to retain existing customers

Data mining can also derive the credit behaviour of individual borrowers with instalment, mortgage and credit card loans, using characteristics such as credit history, length of employment and length of residency.

by offering incentives that are individually tailored to each customer's needs.

When Chase Manhattan Bank in New York began to lose customers to competitors, it began using data mining to analyse customer accounts and make changes in its account requirements, thereby allowing the bank to retain its profitable customers. Data mining is also being used by Fleet Bank, Boston, to identify the best candidates for mutual fund offerings. The bank mines customer demographics and account data along different product lines to determine which customers may be likely to invest in a mutual fund, and this information is used to target those customers. Bank of America's West Coast customer service call centre has its representatives ready with customer profiles gathered from data mining to pitch new products and services that are the most relevant to each individual caller.

Mortgage bankers are also concerned with retaining customers. The program uses leading-edge Internet technologies, predictive models, and customer-direct marketing to enable lenders to identify new customers and retain those that they already have.

DM in Retail Industry

The retail industry is also realising that it is possible to gain a competitive advantage utilising data mining. Retailers have been collecting enormous amounts of data throughout the years, just like the banking industry, and now have the tool needed to sort through this data and find useful pieces of information. For retailers, data mining can be used to provide information on product sales trends, customer buying habits and preferences, supplier lead times and delivery performance, seasonal variations, customer peak traffic periods, and similar predictive data for making proactive decisions. Here are some examples of how the retail industry has been utilising data mining effectively.

Marketing: One of the most widely used areas of data mining for the retail industry, as in the banking industry, is marketing. 'Market basket analysis' is a marketing method used by many retailers to determine optimal locations to promote products. Simply stated, it is the study of retail stock movement data recorded at a Point of Sale (PoS)—to support decisions on shelf-space allocation, store layout, product location and promotion effectiveness. In fact, it uses information about products already purchased by customers to predict which products they would be likely to buy if given special offers or even if they are just made aware of the products. Knowing where to locate products and promote them effectively can increase store sales.

Another marketing tactic employed by many retail stores is the use of 'loyalty' cards. Rewarding customers who are frequent buyers encourages them to do even more of their shopping at that store, and make them less likely to buy from competing stores.

Coupon printers at checkout stands of supermarkets provide an additional way to target customers. These printers are beneficial to brand managers who may not know which customers to target for their brand of products. The coupon printer at the checkout stand can be programmed to print out a coupon for their particular brand when certain products are purchased.

Risk Management: Risk management is another area where data mining is used in the retail industry. However, not as much research has been done in this area as in other areas. Retail organisations use data mining to understand which products may be vulnerable to competitive offers or changing customer-purchasing patterns. Previous purchasing patterns of customers are analysed to identify those customers with low product or brand

A majority of banks in developing countries (particularly in the public sector) are not usually known to exploit their information 'asset' for deriving business value through data mining and gain competitive advantage. But with progressive liberalisation of rules on entry for private and foreign multinational banks, under the GATS framework of WTO, competitive pressure on domestic banks is increasing.

loyalty. Data mining enables retailers to remain competitive and reduce risks by helping them understand what their customers are really doing. Retailers can then target those customers who are more likely to buy a certain brand or product and also be able to promote products in stores where and when they are needed.

Fraud Detection: Retail industries must also be aware that fraud detection is absolutely necessary. Fraud occurring at PoS terminals is a concern for retailers but this can be reduced by using data mining. It is estimated that 38%

of retail shrink occurs because of dishonest employees. And with about 25 paise of every shrink Rupee traceable to PoS fraud, it is no wonder that retailers continue to look for ways to reduce the number of dishonest cashiers.

Some supermarkets have begun to use digitised closed-circuit television (CCTV) systems, along with PoS data mining, to enable retail loss prevention managers to expose cashier stealing and sweet-hearting, assemble convincing evidence, and deal with these situations as a matter of routine. The managers decide what constitutes suspicious behaviour and sends software to detect it. This is called 'exception-based reporting'. The system flags PoS transactions that are the most susceptible to fraud — refunds, credits, discounts, no-sale rings, and the like — and compiles them in a report that identifies the date, time, and checkout lanes where they took place. Managers can then look at the videotape to see exactly what happened; they do not even have to be anywhere near the store.

Customer Acquisition and Retention: Data mining can also help in acquiring and retaining customers in the retail industry. The retail industry deals with high levels of competition, and can use data mining to better understand customers' needs. Retailer can study customers' past purchasing histories and know with what kinds of promotions and incentives to target customers. Also, if a store has seen a number of people leave and go to competitors, data mining can be used to study their past purchasing histories, and use this information to keep other customers from doing likewise.

Conclusion

Data mining is a tool used to extract important information from existing data and enable better decision-making throughout the banking and retail industries. They use data warehousing to combine various data from databases into an acceptable format so that the data can be mined. The data is then analysed and the information that is captured

is used throughout the organisation to support decision-making.

It is universally accepted that many industries (including banking, retail and telecom) are using data mining effectively. Undoubtedly, data mining has many uses in industries. Its practical applications in such areas as analysing medical outcomes, detecting credit card fraud, predicting customer purchase behaviour, predicting the personal interests of Web users, optimising manufacturing processes etc. have been very successful. It has also led to a set of fascinating scientific questions about how computers might automatically learn from past experience. The retail industry is also realising that data mining could give them a competitive advantage.

A majority of the banks in developing countries (particularly in the public sector) are not usually known to exploit their information

Data mining typically involves the use of predictive modelling, forecasting and descriptive modelling techniques. By using these techniques, an organisation can proactively manage customer retention, identify cross-sell and up-sell opportunities, profile and segment customers, set optimal pricing policies, and objectively measure and rank which suppliers are best suited for their needs.

'asset' for deriving business value through data mining and gain competitive advantage. But with progressive liberalisation of rules on entry for private and foreign multinational banks, under the GATS framework of WTO, competitive pressure on domestic banks is increasing. Thus, customer retention and acquisition will be an important determinant of the banks' bottom lines. Those banks and retailers that have realised the utility of data mining and are in the process of building a data mining environment for their decision-making process will reap immense benefit and derive considerable competitive advantage to withstand competition in future. □