

Association for Information Systems

**AIS Electronic Library (AISeL)**

---

ICEB 2003 Proceedings

International Conference on Electronic Business  
(ICEB)

---

Winter 12-9-2003

## **Case Study of Data Mining Application in Banking Industry**

Yongping Liu

Follow this and additional works at: <https://aisel.aisnet.org/iceb2003>

---

This material is brought to you by the International Conference on Electronic Business (ICEB) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICEB 2003 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact [elibrary@aisnet.org](mailto:elibrary@aisnet.org).

# Case Study of Data Mining Application in Banking Industry

Yongping Liu

Applied Mathematics Department  
South China University of Technology  
maypliu@scut.edu.cn

## Abstract

In this paper, we study the usages of data mining in banking industry and its related impacts. Great changes in banking services emerged from the application of data mining especially in retailing banking. We present China Merchant Bank (CMB) as an example to do case analysis, in which we explore data environment evaluation analysis model, operational efficiency model and profitability model to analysis the application performance for CMB. Finally we provide some advices of future development to CMB.

## 1. Introduction

Data warehouse (DW) is like a box, in which vast of data are included and processed into useful information by using various kinds of tools, such as data mining (DM), OLAP, ERP. Banking industry is the pioneer who adopts DW as tool in decision-making. DW makes it possible for business to store large amounts of disparate data in one location. DW consolidates many types of data from many data sources in order to facilitate data analysis for fact-based decision-making.

DW expanded dramatically in the late 1980s, when businesses began realizing the value of their data. Before then, the focus was more on how to integrate DW into a company's existing IT infrastructure than it was about business value. The real value of a DW is to make new business knowledge available literally at the touch of a button. Finding, collecting, and synchronizing data is all performed electronically. Thus, the e-data way of information retrieval provides the needed information in a direct, rapid, and meaningful way, in which queries and answers can be sent and retrieved directly through users and DW. Business intelligence is one of those monikers that encompass the set of products and services for accessing and analyzing data and turning it into

information. In early days DW was motivated by a number of separate factors:

1) Mainframes offload. Frequently, a company's mainframe system was busy enough already, and upgrading could be just as expensive as buying a brand-new DW. These mainframes were often transactional systems performing mission critical functions such as customer billing or banking deposits. Companies were loath to tax their mainframes any further, and the data redundancy and separate processing DW provided were enticing.

2) Dirty data. Different data from all over the enterprise, from accounts payable in compensation to customer information, was often not only hard to find but impossible to understand and often just plain wrong. The DW provided a single platform for loading this "heterogeneous" data. It also offered a pretext for data cleaning---checking it for quality and accuracy, and reformatting it so that it was comprehensive and useful before it was loaded into the warehouse.

3) Security. In order to protect their mission-critical operational systems [6], most companies limit their access to a few experts. DW offered a most generally accessible environment for frequently sought-after information.

In this paper, we study the usages of DM in banking industry and its related impacts. There are 4 parts. In part one and two we explain some background literatures, such as DM, banking services changes emerged from the uses of DM into banking services and retailing banking. In part three and four, we present China Merchant Bank (CMB)[13] as an example to do case analysis, in which we explore 3 models (data environment analysis model, operational efficiency model and profitability models) to make an analysis for CMB. We mention the customer relationship management and personalization in

marketing. At last we provide some advices of future development to CMB.

## 2. Background literatures

DM is the automatic extraction of hidden predictive information from databases[11]. Like statistics, DM is a technology rather than a business solution. For example, consider a catalog retailer who needs to decide who should receive information about a new product. The information operated on by the DM process is contained in a historical database of previous interactions with customers and the features associated with the customers, such as age, zip code, their responses. DM would use this historical information to build a model of customer behavior that could be used to predict which customers would be likely to respond to the new product. By using this information a marketing manager can select only the customers who are most likely to respond. The operational business software can then feed the results of the decision to the appropriate touch point systems such as call centers, web servers, email systems so that the right customers receive the right offers.

Banking industry is the pioneer to adopt DM and its related technology into application. The products banks provide to customers are financial services. Financial services are different from traditional product-making industries. Products provided by financial services are mostly intangible and customer-oriented. To gain greater profit needs maintaining customers' loyalty, which is regarded as customer satisfaction [8]. There are mainly two reasons for banks to adopted DW urgently and necessarily.

1) More and more payment options are available to customers.

New payment technologies and specifications [5] are emerging and need to gain both customers and business acceptance. The ultimate benefits include reduced processing cost, reduced lead times, increased interoperability, and increased flexibility in choosing a payment mechanism. As with any emerging technologies, some will become mainstreams and other products will wither on the vine. Therefore to gain the advantages listed above, banks should equip themselves with solid DW and related technologies to accomplished the task.

2) Rapidly changing customer requirements

The information age is characterized by rapidly changing customer requirements, market niches and focused, global competitors. The information age has greatly influenced strategies business decisions processes, SCM, CRM. The environmental changes experienced by firms in the past decade have been drastic and unpredictable. The innovation of information technology makes huge amounts of customer, market and product data available and accessible to knowledge workers [3].

## 3. Case Analysis of CMB

### 3.1 General history of CMB

CMB was established in 1987 in Shenzhen. With current assets over US\$50 billion and with US\$43.5 billion of deposits, it ranks as the 6<sup>th</sup> largest bank in China. CMB earns good reputation in domestic and aboard markets. Bankers' journal ranked CMB 187th in the top 1000 biggest banks in the world in 2002. CMB was the first bank with public shareholder ownership in China. As the pioneer in China's banking reform, CMB was the first Chinese bank to enter offshore banking and has established international banking agreements with 571 banks in 40 countries worldwide. Within China, CMB has 355 branches with its own nation-wide commercial banking network. CMB is also the first China bank to officially establish investment-banking operations in the Chinese banking industry, and more recently, has set up the first China Internet banking service. CMB was also the first bank in China to obtain ISO9001 certification.

### 3.2 Technology advantages of CMB

During 16 years development, CMB has been consistently adopting the technology driven strategy. In 1995 CMB put all-in-one credit card into market, which was regarded as a milestone of personal financial management in China banking industry. In the next year, CMB was the first bank to turn nationally credit and debit into reality, while in the same year, it launched IC card, which makes it possible to withdraw and deposit without distance and time limitation among companies. In addition, CMB also built the website of national ATM and domestic consumptions of POS, both of which

compose modernized and national personal financial service network supported by powerful and advanced technology, the volume of all-in-one credit card has been drastically increase and becomes the shining product of CMB. In Feb 1998, all-in-net service is developed by CMB, which is the first bank to launch on-line banking service in China. 21 months later CMB became perfect on-line corporation banks, on-line personal banks, on-line trading mall and on-line security investment and on-line payment. Till now, CMB's on-line banking is occupying an absolutely leading place in technology domain and service income area in banking sector of China. About 80% famous web sites, for instance, www.sohu.com, www.sina.com.cn , www.8848.com, choose "all-in-net" as the first or even the unique choice as payment tool. The achievement CMB has gained is not accidental. It is due to its strong and solid DW system. Moreover, In 2000 China Merchant Bank joined the ESS Technology corp., a leader in Internet, communications, PC audio and digital video, and China E-Commerce Technology Co. Ltd. (CECT) and announces a major breakthrough in Chinese Internet e-Commerce by creating the first secure e-Commerce payment gateway in China. The three companies jointly announced the new solution to facilitate payments through an Internet Set-Top-Box designed by ESS using ESS's technologies and products. The solution utilizes ESS's Internet Set-Top-Box to protect both the account holders and the bank. The Set-Top-Box automatically connects to CECT's portal and ICP website that operates the CMB gateway. CMB serves as the clearing bank, which settles all e-Commerce transactions through a debit system for each customer's account.

### 3.3 Why CMB use DW and DM

The DW is a site, which combines information from different, separate systems in one location. Because the information is stored together, it is easy to access. Data in the warehouse comes from a number of sources. At this point, the warehouse of CMB contains four main types of information identified by the business area: financial, general, individuals and corporations. The DW is a read-only database, which guarantees its stability over time. To do customer segmentation, CMB needs DM

urgently. For example, a pattern might indicate that married males with children are twice as likely to drive a particular sports car than married males with no children. If you are a marketing manager for an auto manufacturer, this somewhat surprising pattern might be quite valuable. DM uses well-established statistical and machine learning techniques to build models that predict customer behavior. The technology enhances the procedure by automating the mining process, integrating it with commercial DW, and presenting it in a relevant way for business users. DM helps banks improve their understanding of customer behavior so as to target marketing campaigns more accurately and to align campaigns more closely with the needs, wants and attitudes of customers and prospects. The key is to find patterns relevant to current business problems.

### 3.4 Principles of DM application in CMB

According to a report that released by CMB in 2002, there are several thoughts on the current state of DM software applications in CMB:

#### 1) Automated Model Scoring

Scoring models against a database is currently a time consuming, error prone activity that hasn't been given enough consideration [2]. When someone in marketing needs to have a database scored, they usually have to call someone in IT and cross their fingers that it will be done correctly. If the marketing campaigns that rely on the scores are run on daily basis, this means a lot of phone calls and lot of manual processing. Instead, the process that makes use of the scores should drive the model scoring. Scoring should be integrated with the driving applications via published API and run-time-library scoring engines. Automation will reduce processing time, allow for the most up-to-date data to be used, and reduce error.

#### 2) Time-Series Data

Much of the data that exists in DW has a time-based component. A year's worth of monthly balance information is qualitatively different than twelve distinct non-time-series variables. DM applications need to understand that fact and use it to create better models. Knowing that a set of variables is a time-series allows for calculations to be done that make sense only for time

series data: trends, slopes, etc.

### 3) Incorporate Financial Information

The results of the DM process will drive efforts in areas such as marketing, risk management, and credit scoring. Each of these areas is influenced by financial considerations that need to be incorporated in the DM modeling process. The information necessary to make these financial decisions is often available and should be provided as an input to the DM application.

### 4) E-banking and personalized services

CMB retain credibility by being able to offer Internet access and defend existing customer base against potential competitors. The e-banking and personalized services are seemed as something that CMB cannot ignore if it wants to fight out a way in banking industries. The Internet is particular well suited to the distribution of financial products. Fortunately there are tangible opportunities to shift manual transactions out of branches and thereby reduce costs the technology will continue to improve.

Now we are going to analysis 2 kinds of the knockout products CMB developed when combining the DW technique.

#### 1) All-in-one credit card

Since July 1995, all-in-one credit card created a new age of personal financial management depending on its advanced technology and fashionable business concept. It has many functions, some of which are also included in credit cards other banks issue, such as arranged transactions, automatic transfers, customer demands, ATM. But the functionalities such as mobile service banking, security, automatic payment, transfers and transactions, montage, pay on net, long distance call, hotel reservation, are characterized with personal financial management services in which other banks are deficient. With the help of the DW technology, CMB is able to provide the services to the customers that tailored to their particular needs. Related card is one of the typical examples. Related card is produced by Nanjing branch and Nanjing Security bureau, which is added with some extra functions especially for the drivers. According to different individual drivers, inquiry of transportation management information and pay traffic fines are accessible in related cards.

#### 2) Mobile banking service (MBS)

CMB cooperated with China mobile corp. and put MBS into marketing. This is a newborn personal banking service in which telecommunication techniques are utilized in order to make personal financial management more efficient than before. MBS contains many personal investment functions, from the common ones like account query, payments, functions to multifunction deposit forms transferring, moreover it also includes advanced personal investment function, as customers can do the daily security information inquiry, entrusting transactions, international currency exchange and inquiry through mobile phone.

## 4. Models to analysis for CMB

Bank management and government regulators need models that can accurately portray value creation in banking and offer guidelines for strategy and the regulation of strategic behavior by banks. Internet and telephone banking usage per person in the world is increasing rapidly, thereby underlining that banks are intermediaries facilitating direct and indirect financial relations among their customers, rather than creators and distributors of financial products.

### 4.1 Data Envelopment Analysis (DEA)

DEA is receiving increasing importance as a tool for evaluating and improving the performance of manufacturing and service operations [10][12]. It has been extensively applied in performance evaluation and benchmarking of bank branches. We make an introduction to the DEA and some importance methodological extensions that have improved its effectiveness as a productivity analysis tool for CMB branches. DEA is a multi-factor productivity analysis model for measuring the relative efficiencies of a homogenous set of the decision making units (DMUs). The efficiency score in the presence of multiple input and output factors is defined as

$$\text{Efficiency} = \frac{\text{weighted sum of outputs}}{\text{weighted sum of inputs}} \quad (1)$$

Assuming that there are  $n$  DMUs, each with  $m$  input and  $s$  outputs, the relative efficiency score of a test DMU  $p$  is obtained by solving the following model proposed by

Charnes et al.:

$$\begin{aligned} & \max \frac{\sum_{k=1}^s v_k y_{kp}}{\sum_{j=1}^m u_j x_{jp}} \\ \text{s.t. } & \sum_{k=1}^s v_k y_{ki} \leq \sum_{j=1}^m u_j x_{ji}, \quad \forall i = 1 \dots n \\ & v_k, u_j \geq 0, \quad \forall k = 1 \dots s, \quad j = 1 \dots m \end{aligned} \quad (2)$$

$x_{ji}$  = Amount of input  $j$  utilized by DMUs  $i$ ,  
 $y_{ki}$  = Amount of output  $k$  produced by DMUs  $i$   
 $u_j$  = Weight given to input  $j$ ,  
 $v_k$  = Weight given to output  $k$

The fractional program shown as (2) can be converted to a linear program as shown in (3).

$$\begin{aligned} & \max \sum_{k=1}^s v_k y_{kp} \\ \text{s.t. } & \sum_{j=1}^m u_j x_{jp} = 1 \\ & \sum_{k=1}^s v_k y_{ki} - \sum_{j=1}^m u_j x_{ji} \leq 0 \quad \forall i = 1 \dots n \\ & v_k, u_j \geq 0 \quad \forall k = 1 \dots s, \quad j = 1 \dots m \end{aligned} \quad (3)$$

The above problem is run  $n$  times in identify the relative efficiency scores of all the DMUs. Each DMU selects input and output weights that maximize its efficiency score. In general, a DMU is considered to be efficient if it obtains a score of 1. A score of less than 1 implies that it is inefficient. The primary advantages of this technique are that it considers multiple factors and does not require parametric assumptions of traditional multivariate methods. However there are some critical factors one must consider in the application of DEA models. The discovery scores could be very sensitive to changes in the data and depend heavily on the number and type of input and output factors considered. In general, inputs can include any recourses utilities by a DMU, and the output can range from actual products produced to a range of performance and activity measures. The size of the data set is also an important factor when using some of the traditional DEA models.

## 4.2 Operational efficiency model

Broadly speaking, operational efficiency model adopt either a production approach or an intermediation approach [4]. In the former case the branch is considered as a factory delivering services to its clients in the form of transactions. Benchmarking models examine how well different branches combine their resources to support the largest possible number of transactions. The intermediation approach considers various types of costs as the inputs, and those are combined to support the largest possible number of revenue generating accounts.

### 1) Model input

The model uses two broad sets of input. One set captures the resources used by the branch. The second set of inputs includes the number of accounts in different account categories. The clientele structure is tightly linked to a specific branch and it changes very slowly with time. Hence, for the purpose of a static analysis this information is part of the operating environment of the branch.

### 2) Model outputs

The outputs of the model is total amount of work produces by the branch in order to support the given client base. This work is produced in the form of transactions carried out of the various clients and their accounts. A standardizes measure of total work can be obtained by counting the total number of elementary tasks required in order to be executed for these transactions during a working day. Each elementary task requires summing up their standard times in order to be completed. The standard times of each task, for example, filling in a form or getting a supervisor's signature is obtained through a work measured as the total time involved in processing the tasks for all transactions carried out at a branch.

## 4.3 Profitability efficiency model

We now turn to the final link of the service-profit chain and address the issue of profitability efficiency. Substantial research has been done on this key issue [1][2]. We formulate the model of benchmarking the profitability.

### 1) Model inputs

The inputs are the resources used by the branch and revenue generating accounts. However that for some

types of accounts is indicative of the demands imposed on the branch's resources for service and maintenance, while the volume of deposits in these accounts is indicative of their contribution to profits. Hence, the accounts of each branched are measured using a different classification and different units than those used in the previous model.

#### 2) Model outputs

The outputs are the profits generated at each branch. Revenue growth could be incorporated as a more appropriate output, but then other inputs should also be included such as service quality, demographics of the region where the branch is located, and competition. The model can be readily extended if such data is available.

### **4.4 Linking the operational efficiency with profitability**

We use operational efficiency model and the profitability efficiency model to analysis the effects of inefficient operations and the costs of offering inappropriate products at each branch. The profitability efficiency model considers both consumable and revenue generating resources as inputs. The consumable resource is operational variables whereas the revenue generating resource is the product of each branch.

### **5. Retail Banking of CMB and Customers Relationship Management**

The industry of universal banking is restructuring, and the customer has never met a richer supply of information and services. The old relationship between bank and customer is changing. Individual consumers are becoming ever more demanding, both on kinds and quality of services. CMB are looking to gain a competitive edge by establishing one-to-one relationships with customers and a higher degree of personalization. For example the profitability of various channels can be notoriously difficult to ascertain accurately, but the profitability of individuals using a particular ATM could be established very easily, aiding decision on whether that equipment was worth retaining or upgrading. Similarly the data might indicate that a particular product is being used by a large number of customers in a very unprofitable way. Also by using DW technology, CMB is able to do customer segmentation:

1) CMB can map the results of customer profitability analyses back to their channels to determined high-value branches, or to specific geographical areas to extrapolate high-value cities or neighborhoods to target. Changes in a customer's profitability over time can also uncover previously unknown customer behavior patterns. Many banks use profitability analysis to the creation of new services, or discarding of old ones. CMB is likely to target particular customer segment, e.g. mass affluent rather than mass market and will try to differentiate itself from existing competitors.

#### 2) Increasing customer lifetime value

Consider, for example, customers of CMB who only use the CMB for a checking account. An analysis reveals that after depositing large annual income bonuses, some customers wait for their funds to clear before moving the money quickly into their stock-brokerage or mutual fund accounts outside the bank. This represents a loss of business for the CMB. To persuade these customers to keep their money in the bank, CMB can use DM tools to immediately identify large deposits and trigger a response. The system might automatically schedule a direct mail or telemarketing promotion as soon as a customer's balance exceeds a predetermined amount. Based on the size of the deposit, the triggered promotion can then provide an appropriate incentive that encourages customers to invest their money in the bank's other products. Finally, by tracking responses and following rules for attributing customer behavior, DM tools can help measure the profitability and ROI of all ongoing campaigns.

### **6. Personalizations and Customer Oriented Marketing**

The Internet is leading marketer to a fundamental paradigm shift from mass marketing to personalized marketing. Databases, cookies and telecommunications technology make it very easy and cost-efficient to mass-market personalized services. Personalization on the Internet refers to the ability of customers to receive personalized information or visit a website with a home page customized for them. Through the use of databases, it is able to store information about its visitors and provide a personalized, free service for them: remaining individuals of important events, such as birthdays,

anniversaries, and graduations. Personalization represents simultaneously the extreme challenge and opportunity present in the form of Internet medium. A fundamental shift is occurring between producer and consumer. Technology-enabled consumers are actively participating in activities, tasks and decisions that were once the sole domain of the marketer. In the past, consumers' decisions were primarily limited to purchase and consumption. Today through the enabling power of technology, consumers are able to design products from their homes, and store and access information at their convenience [7].

## 7. What Should CMB Improve in Personal Service Area?

### 1) Financial Advice

One higher level of CMB goes beyond typical to offer more general financial advice, sometimes with a level of interactivity [9]. An important decision that banks will face in the future is how to develop their personalities and how to position themselves in their non-banking capacities. This means understanding customers better and becoming the sort of bank that they would want to do business with.

### 2) Selling Financial Products

Some banks have gone so far as to sell financial products on the Internet. At this stage it is most common to allow customers to fill in an application form for a credit card on-line and complete the transaction by more mundane methods such as the post, but we can expect that the future will see banks moving to complete sales entirely on-line, and to sell more complex products. Questions are raised such as: what will people be willing to buy on-line? What sort of people will want to make purchases? How secure and confidential can the information generated by a sale be? There are many other questions such as these, which for the time being are difficult to answer, but are likely to emerge as central areas of research not only for banks but for the business world in general.

### 3) Privacy

In recent years privacy concerns have taken on a more significant role not only in China but also all over the world as merchants, insurance companies, and government agencies amass warehouses containing

personal data. The concerns that people have over the collection of this data will naturally extend to any analytic capabilities applied to the data. Users of DM should start thinking about how their use of this technology will be impacted by legal issues related to privacy. For example, if the primary purpose of the collection of transactional information is to permit a credit card payment, then using the information for other purposes, such as DM, without having identified this purpose before or at the time of the collection, is in violation of the purpose and use limitation principles. The primary purpose of the collection must be clearly understood by the consumer and identified at the time of the collection. Since DM is based on the extraction of unknown patterns from a database, DM does not know, cannot know, at the outset, what personal data will be of value or what relationships will emerge. Therefore, identifying a primary purpose at the beginning of the process, and then restricting one's use of the data to that purpose are the must.

## 8. Summary

DW empowers banking institutions to provide Internet-based financial solutions to their retail and commercial customers via a cost-effective service. DW offers profit-generating opportunities to its clients through secure, reliable, and scalable products and services. Few business areas are changing so rapidly as the financial services area. Banks have a long history of conservatism in their business practices. However, for the past decade, changes in technology, deregulation and general economic conditions have transformed the industry. The DW technology is rapidly becoming the new medium for financial services. The ramifications of the changes are unpredictable. Many financial services firms outsource their software and hardware development to computer companies. The DW technology makes it possible for large global financial services firms to reach the most remote local markets, but also for small specialized niche banks to reach their customers world wide.

## Reference

[1] A.J.Zahorik, T.L. Keiningham. Return on quality (ROQ): making services quality financially accountable. *J.Marketing*, 1995, 5(9), 58-70.



- [2] Andreas Soterious, Stavros A. Zenios, Operation, quality, and profitability in the provision of banking services, Prentice-Hall, 2000.
- [3] Barua, Information technologies and business value: an analytic and empirical investigation, *Information Systems Reserch*, 1995, 6(1), 3-23.
- [4] Berger, A.N., L.J.Mester. Inside the black box: what explains differencies in the efficiencies of financial institutions. *J. Banking and Finance*, 1997, 4(70), 859-947.
- [5] Bowen, J.W., Hedges, R.B., Increasing services quality in retail banking. *J. Retail banking* 1993, 15(3), 21-28.
- [6] Marilyn Greenstein, Todd M.Feinman, Electronic commerce: security, risk, management and control, McGraw-Hill/Irwin, 2001.
- [7] Michael D. Johnson, Anders Gustafsson, Improving customer satisfaction, loyalty, and profit: an integrated measurement and management system, Jossey-Bass, 2000.
- [8] Naumann, E., K. Giel. Customer satisfaction measurement and management. Thomson Executive Press, 1995.
- [9] Peterson, R. A. Measuring customer satisfaction: facts and artifacts. *J. Marketing Sci.* 2000, 20(1), 61-71
- [10] P.T. Harker. Projections onto efficient frontiers: theoretical and computational extension to DEA. *J. Productivity Anal.* 1999, 11(3), 275-300.
- [11] R. Groth, Data mining: building competitve advantages, Prentice-Hall, 2001.
- [12] William W. Cooper, Lawrence M. Seiford, Kaoru Tone, Data envelopment analysis: a comprehensive text with models, applications, references and dea-solver software, Kluwer Academic Pub, 1999.
- [13] WWW.CMBCHINA.COM