

# Application of Data Mining in the Banking Sector

Dilmuradjon ZAKIROV \*  
Nodar MOMTSELIDZE \*\*

## Abstract

Data mining is one of the important functions of information communication technology for business. Data mining refers to the process of collecting data and analyzing it for/with the aim of achieving a specific goal using the data. In the research, the data from secondary resources was used to create the article. From these sources, it is evident that data mining as a modern method of analyzing data is more appropriate compared to the traditional method. The accuracy of data analysis is also very accurate compared to the traditional approach. In business organizations, it is evident that data mining is being used in the banking area in: managing risks, doing marketing, mobile banking, offering customer relationship management and fraud detection.

**Keywords:** data mining, fraud detection, mobile banking, supervised and unsupervised learning

## Introduction

Information and communication technologies (ICT) are perceived as one of the handles of the modern society. This is due to the fact that it is one of the most used applications in business, governmental organizations and other forms of organizations. Data mining is one of the aspects associated with the usage of information technology. Data mining refers to the process of discovering interesting information from a huge amount of data or information (Zhao, 2013, p. 1). The data solicited through data mining should be capable to serve the intended purpose. Data mining technique involves various activities such as sequence analysis, clustering, prediction, classification, association rules, sentiment and social network analysis, text mining, outlier detection and time series analysis. These aspects are some of the important areas of bioinformatics, finance, telecommunication, retail and even banking.

## Method

The method involved in this research was basically secondary research. It involved researching through the peer reviewed secondary sources regarding data mining. Various information regarding data mining was recorded and analyzed to provide a wide understanding of data mining concepts.

## Findings

It was evident from the results that data mining is a com-

plex concept and it involves various techniques. These techniques and concepts are explained under the various sections in the article.

## Types of data mining

There are basically two categories of data mining – supervised and unsupervised learning. Concepts of both categories are derived from machine learning – a technique “that enable devices to learn from their own performance and modify their own functioning” (Chitra & Subashini, 2013, p. 220).

The process of data mining involves the application of these concepts from machine learning to the data. Supervised learning is also referred as “directed learning”, which means to say that “the learning process is directed by a previously known dependent attribute or target” (Chitra & Subashini, 2013, p. 220). The goal of this category is to explain the target’s behavior in regard to its set function of independent predictor or attribute. This is contrary to the goal of the unsupervised learning – pattern detection. Unsupervised learning category is non-directed learning (Chitra & Subashini, 2013, p. 221).

## Steps and techniques used in data mining

It is important to discuss the different steps and techniques of data mining to see “how unprocessed data converts into

\* Head of Information Technologies Department, Demir Kyrgyz International Bank, Bishkek, Kyrgyzstan  
E-mail: dilmuratz@gmail.com

\*\* Prof.Dr., Faculty of Computer Technologies and Engineering, International Black Sea University, Tbilisi, Georgia  
E-mail: nmomtselidze@ibsu.edu.ge

meaningful information” (Jayasree & Balan, 2013, p. 1161):

1. Data Selection – this step involves identifying the data location and its importance to the objective of the institution. Due to the pervasive characteristic of electronic data, it is important to take into consideration its quality.

2. Data Preparation – after identifying the location, the data will undergo integration and cleaning. Cleaning includes the process of removing irrelevant and noise data while integration involves the clustering of the same data sources. Afterwards, the data are set to analyze. Data analyzing may come in two forms based on the accessibility of the appropriate and existing data source models – confirmatory or exploratory. Nevertheless, these models are geared towards “grouping, clustering or classification of measurements based on their goodness of fit to a postulated model or natural groupings (clustering) revealed through analysis” (Jayasree & Balan, 2013, p. 1161).

3. Data Mining – this is the most important step as the selected data will be converted into the suitable form for the data input.

4. Evaluation – this process involves the identification of patterns that represent certain knowledge based on the given measures.

5. Representation – this step involves the visual representation of the newly discovered knowledge. This is based on the assumption that visualization of techniques are deemed effective to understand the end users’ output.

### Traditional approach vs data mining model in the banking sector

Traditional method was the one that had been used before the advent of data mining model. This traditional approach is very time consuming and complex, thus making the investigation slow and in most cases unsuccessful. Data mining is deemed as the solution to this problem due to the fact that the approach is more efficient and relatively accurate compared to the traditional method. Through the application of data mining, banks can easily come up with models that can be used to predict various operations of the business. These systems can as well visualize the actual report and discern meaningful information from the data to the users. There are various models regarding the application of data mining in the banks. This is illustrated in the following figure. Figure 1 is one of those models proposed by Chitra & Subashani (2013).

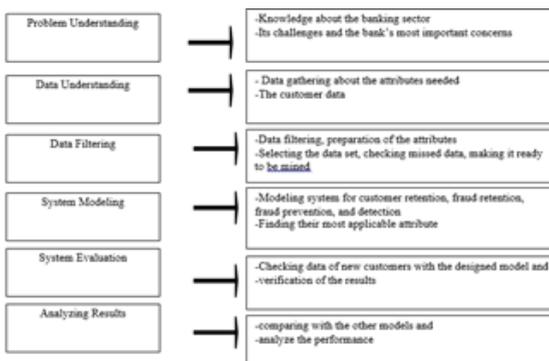


Figure 1. Data Mining Techniques and Processes

Through data mining, large databases could be predict-

ed automatically for future behaviors and trends. Data could also be discovered automatically for the previously unknown patterns. One of the industries that benefited from the development of information and technologies is the banking sector (Pulakkazhy & Balan, 2013, p. 1252). For instance, prior to the introduction of data mining, banks stored data in their respective branches. Today, these branches enjoy the comfort of storing large data in a centralized location. Customers also enjoy the efficiency of online transactions, such as ATM, electronic transfers, and depositing checks and cash. The figures below compare and contrast the differences between the traditional way of decision making and the application of data mining in the banking sector. The figures were adopted from Pulakkazhy & Balan, 2013, p. 1253.

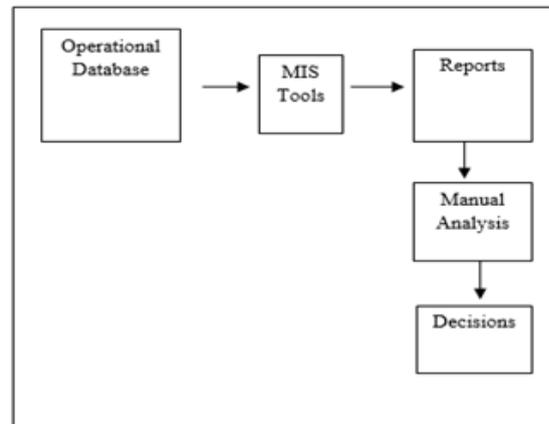


Figure 2. Traditional Decision Making Process

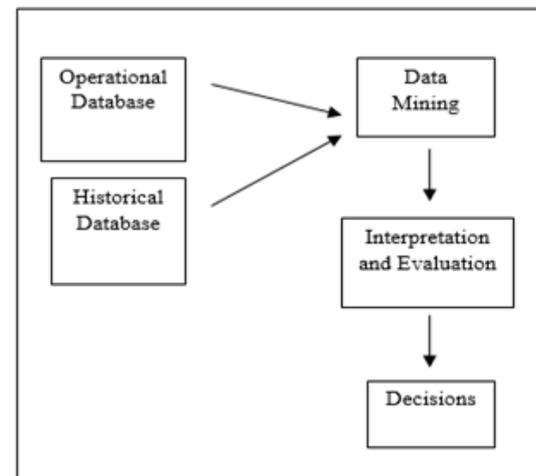


Figure 3. Decision Making with Data Mining

In the figure above, it can be observed that the process of decision-making takes longer in the traditional model. Further, the data volume to analyze is limited because of the capacity of the tools used. Therefore, the decision that may be extracted from the analyzed information may not be efficient and accurate. For instance, “it could be possible that loan installments are being paid regularly though there is an alarming negative trend in the customers turnover and the account may be about to default” (Pulakkazhy & Balan, 2013, p. 1253). This example could not be easily detected if the bank uses traditional process. On the other hand, figure 3 shows how efficient it is to come up with a decision through the use of data mining. Large set of data, including historical base, could be interpreted and analyzed.

Application of Data Mining in the Banking Sector In the banking sector, there are several applications of data mining – credit analysis, cross selling, customer profitability and segmentation, fraudulent transactions, ranking investments, most profitable customers on cross selling and credit card, and the like. Due to the growing competency in the market, it is highly encouraged that banks analyze customer profiles and preferences in order to retain them (Pulakkhazhy & Balan, 2013, p. 1252). It is important to analyze transaction patterns in order to track whether or not fraud transactions will directly affect the profitability of the bank.

## Risk Management

Another application of data mining in the banking sector involves default detection and risk management. Risk management is one of the very important phrases in project management these days. The fundamental concept of risk management is a straightforward evaluation of the project which recognizes anything important that could probably go wrong with the project. For small businesses, uncomplicated and practical techniques can be followed namely risk identification, risk analysis, risk response and risk monitoring and control. In the risk identification, the company starts with knowing that something will possibly go wrong and unanticipated things might happen. The main focus in the risk identification is to detect risks in a positive way. In short, the project managers should figure out what could go wrong and list it down.

Quantifying the risk of lending decision does not only limit the financial loss' risk in the bank but also make the process of risk management easier (Pulakkhazhy & Balan, 2013, p. 1256). Information about the capability of a specific customer to repay can be detected through the use of data mining. This information can aid the credit manager in his/her decision. Further, through data mining, banking personnel can identify customers who will default or delay their loan payment. In this way, banks can come up with the decision before it is too late. They can make corrective measures in order to prevent possible losses. They can consider limit utilization, check return, turnover trends, and behavioral patterns. This information can also be used to predict future patterns (Pulakkhazhy & Balan, 2013, p. 1256).

Through data mining, bank can also secure information about the customer's credit and behavior scores. This information is deemed important in the process of risk management in the banking sector. Credit score denotes the creditworthiness of the borrower while the latter pertains to the behavior of the customer, which can be obtained from the probability models. Behavioral score is important in predicting the potential future behavior of a certain customer. For instance, data mining can evaluate the banking profile of a customer who wants to avail bank loan. Through appropriate data mining techniques, the bank can see whether the candidate has a good credit history or high salary. By having this information, the bank can reduce the default probability of the loan customers.

## Marketing

Another application of data mining in the banking sector can be seen in the marketing aspect. Data mining techniques provide trend analysis on the unique products with high quality. Further, the realm of marketing, the banking sector can also cluster their clients according to their performance: "highly potential, good, low" (Jayasree & Balan, 2013, p. 1163). By classifying their customers, banks can provide appropriate

products and services in accordance to their clusters. Also, data mining can trace the reaction of customers whenever the bank makes adjustment in the installment changes and interest rates on borrowing and depositing products. Data mining can also trace clients who are active in smart phones or web applications of the banks. Hence, this can aid the organization to develop further their electronic and telemarketing applications.

## Mobile Banking

Another application of data mining is the efficient application of mobile banking. Mobile banking is defined "as carrying out banking business with the help of mobile devices such as mobile phones or PDAs" (Tiwari, Buse, Herstatt 2006, p 6). Through mobile banking, banks are able to expand beyond their geographical location. Thus, they can up-sell and cross-sell their products to their present and potential customers. One microfinance bank, the Opportunity Bank in Malawi, had proven the effectiveness of the use of technological innovations in the bank operations. In the year 2007, the bank was able to maximize its customers (reaching 150,000) despite of the fact that the area has poor infrastructures and more than half of the population lives in the rural area (Kumar, McKay and Rotman 2010, p. 1). This condition implies that reaching customers and potential customers is difficult to achieve.

The chief executive officer of Opportunity Bank, Aleksandr Kalanda, believed that the success of their branch should not have possibly met if not because of technological innovations like the introduction of mobile banking. This innovation is "the delivery of financial services outside the conventional bank branches using mobile phones and non-bank retail agents" (Kumar, McKay, and Rotman 2010, p. 1). Because of its ability to deliver services regardless of where the customer is, scholars see its potential to serve better the existing customers and reach more potential customers. It is also worth to note the difference between mobile banking and mobile payments. While mobile banking entails the platform that allows customers to enjoy financial services such as investment options, bill payments, transfers and balance information, mobile payments "is generally defined as the process of using a hand-held device to pay for a product or service, either remotely or at a point-of sale" (KPMG 2011, p. 4).

## Customer Relationship Management

CRM is a mixture of hardware, software, procedure, relevance and administration assurance to develop client service, maintain clients and present methodical capabilities (Roberts & Wallace 2002, p. 14). The vibrant practice of running a Customer-company connection makes the purchasers prolong mutually advantageous profit-making exchanges and discourage from involving in exchanges that are unbeneficial to the business (Smith 2009, p. 4). As the competition in market gets tighter, customer loyalty has become one of the target marketing objectives of CRM. Reaching customers is one of the most challenging problems that microfinance institutions face. Many organizations are trying to experiment and design different alternative means that are not too costly and at the same time will reach customers and maximize their convenience.

Chitra & Subashani (2013) claimed that data mining is used to improve the customer relationship management (CRM) in the banking sector: acquisition, increase value, and retention of customers. Improving CRM would mean

helping banks to “build long-lasting relationships with their customers and increase their revenues and profits” (Chitra & Subashani, 2013, p. 219). This aspect is very important in the banking sector as it goes beyond acquiring customers but customer retention. Through extracting information about the factors that affect the buying behavior of a customer, the banking sector can improve its marketing strategy (Pulakkazhy & Balan, 2013, p. 1256). Data mining models that can be used in this aspect are behavioral scoring and credit scoring models. In order to improve the bank’s customer retention through data mining, there are three steps that should be followed: 1) customer retention measurement; 2) identifying the underlying factors that affect defection including key issues; 3) designing corrective action in response to the defection.

## Fraud Detection

Another application of data mining is fraud detection. This process involves “the recognition of symptoms of fraud where no prior suspicion or tendency to fraud exists” (Jayasree & Balan, 2013, p.1161). This type of deception is rampant, especially in the banking sector, stock market, mobile companies, and insurance agencies. There are several data mining techniques that can be used to detect fraud. One of them is the record screening modeling technique (Tpistsis & Chorianopoulos, 2009, p. 59). This technique allows the bank to identify odd and unexpected cases and data patterns. Though this does not automatically indicate a suspicious action, this information will lead to further investigation. Banks can also utilize unsupervised and supervised models in the process of detecting fraud. For instance, if there is available information about past cases of fraudulence, analysts may use supervised model by identifying “the input data patterns associated with the target suspicious activities” (Tpistsis & Chorianopoulos, 2009, p. 60).

## Cross-Selling

Cross-selling is another aspect where data mining could be found useful. This is a process whereby “a service provider makes it attractive for a customer to buy additional products or services with the same business” (The Chartered Accountant, 2006, p. 591). This is based on the assumption that the more services and products a bank gives to the customer, the higher possibility of retaining these customers. Cross-selling allows banks to know about the “purchasing potential” of their existing clients. Cross-selling based on data mining may be divided into two forms – up selling and deep selling. The former pertains to the ability of the banks to offer and switch customers to their premium products that are more expensive (hence more profitable) than their current product. On the other hand, deep selling pertains to the ability of the banks to let their customers increase their usage of the services and products that they currently have (Tpistsis & Chorianopoulos, 2009, p. 33).

## Discussion and Conclusion

We have entered in the world where computers have become essential in the course of human daily activities. In as much as the advent of computer technologies, it has extended its influence not only in the household but also in the realm of the banking sector. The last two decades of the 20th century marked tremendous growth in electronic applications because of the increasingly used information and communication technology (ICT), especially the Internet,

to deliver banking services in all forms. The application of data mining in the banking sector has reached its peak and become popular nowadays. Because of the demand of an easier and faster task, it is prerequisite for the services to be platform-independent and constantly available. Generally, banks use data mining because of its capability to offer the lowest cost in storing data yet increased ease when it comes to data collection. Thus, data mining helps banks to understand their own business, allows the organization to serve better their clients, and improve the organization’s effectiveness. With the help of the Internet, banking services based on data mining has reached maturity as an effective means of reaching out the business sector and the customers.

## References

- Chitra, K. & Subashini, B. (2013). Data Mining Techniques and its Applications in Banking Sector. *International Journal of Emerging Technology and Advanced Engineering* 3 (8): 219-226.
- Jayasree, V. & Balan, R. (2013). A Review on Data Mining in Banking Sector. *American Journal of Applied Sciences* 10 (10): 1160-1165.
- Kumar, K., McKay, C. and Rotman, S. (2010). Microfinance and Mobile Banking: The Story so Far. CGAP Publications.
- KPMG (2011). Monetizing Mobile: How Banks are Preserving their Place in the Payment Value Chain. KPMG International Cooperative.
- Pulakkazhy, S. & Balan, R. (2013). Data Mining in Banking And Its Applications: A Review. *Journal of Computer Science* 9 (10): 1252-1259.
- Roberts, A. & Wallace, W. (2002). *Project Management*. Great Britain: Edinburgh Business School Publications
- Smith, C. (2009). *Facebook Is The Future Of CRM*. CRM Magazine.
- The Chartered Accountant (2006). Data Mining: A Competitive Tool in the Banking and Retail Industries. *Banking and Finance* 1: 588-594.
- Tpistsis, K. & Chorianopoulos, A. (2009). *Data Mining Techniques in CRM: Inside Customer Segmentation*. John Wiley and Sons, Ltd., Publication