

## IMPLEMENTATION OF DATA MINING IN BANKING SECTOR- A FEASIBILITY STUDY

Vivek Bhambri\*

---

### ABSTRACT

*Banking sector around the globe is using information technology for their day to day operations and banks have realized this fact that their biggest asset is the knowledge and not the financial resources. Banks have started using the latest techniques of data mining for customer segmentation, profiling, fraud detection, targeting the new customers, retaining the existing customers and for predicting the future trends. The primary objective of this paper is to check the feasibility of implementation of these techniques in India banking sector. The economic, technical and operational feasibility has been investigated to explore whether the industry is ready for the implementation of these techniques or not. The survey of private and public banks has been done to assess the feasibility of implementation of techniques of data mining.*

**Keywords:** *Data Mining, Economic Feasibility, Fraud Detection, Segmentation.*

---

\*Research Scholar, Singhania University, Pacheri Bari, Jhunjhunu, Rajasthan

## INTRODUCTION

Information and Communication Technologies (ICTs) are now widely accepted as key forces in shaping the economic landscape, transforming the way we live, learn, work and play (Danziger & Andersen, 2002)<sup>1</sup>. The size of data bases is growing at an alarming rate, but the problem is that the companies fail to reap the full benefits which can be gained from this great wealth of information. The banking sector is not an exception. The banking sector has witnessed a tremendous change in the way the banking operations are being carried out. Since 1990's the whole concept of banking has been shifted to centralized databases, online transactions, network connectivity and ATM's all over the world, which has made the banking system technically strong and more customer oriented. Banks have to adjust to the changing needs of the societies, where people not only regard a bank account as a right rather than a privilege, but are also aware of the fact that their business is valuable to the bank, and if the bank does not look after them, they can take their business elsewhere (Engler & Essinger, 2000)<sup>2</sup>. Technology in banking is not just the automation of process, but it is much more than this. Where the age old product-strategy has be transformed to customer-focused strategy. The huge size of data bases makes it impossible for the organizations to analyze it as per the need of the decision makers. So the need of data mining techniques is being felt.

A data mining approach comprises of a variety of technical techniques and tools to explore, summarize, compare, analyze, forecast and estimate the data for various decisions to be taken by the organization. It provides a methodology for problem solving, analysis, planning, diagnosis, detection, integration, prevention, learning and innovations (Hedelin & Allwood, 2002, Liao, 2003)<sup>3</sup>.

The objective of data mining is to identify valid, novel, potentially useful, and understandable correlations and patterns in existing data. Chung and Grey-

Data mining is a process that uses statistics, artificial intelligence and machine learning techniques to extract and identify useful information, and subsequent knowledge, from large databases. Nemati & Barko (2003)<sup>4</sup>

## LITERATURE REVIEW

Keeping the requirement of use of information technology in the banking sector, the Reserve Bank of India constituted a committee on technology up gradation in the banking sector (1999)<sup>5</sup>, the committee highlighted the usage of management information systems by the banks and recommended that by the use of data mining techniques, data available at various computer systems can be accessed and by a combination of techniques like classification,

clustering, segmentation, association rules, sequencing, decision tree various ALM reports such as Statement of Structural Liquidity, Statement of Interest Rate Sensitivity etc. or accounting reports like Balance Sheet and Profit & Loss Account can be generated instantaneously for any desired period/date. Trends can be analyzed and predicted with the availability of historical data and the data warehouse assures that everyone is using the same data at the same level of extraction, which eliminates conflicting analytical results and arguments over the source and quality of data used for analysis. Rob Gerritsen(1999)<sup>6</sup> opines that the performance of the company will improve many folds if they could predict the response of the customers for a particular promotion scheme or if the company could predict that whether the customers will be able to repay the loans or not. The basic data mining techniques and models were applied in a project for the US Department of Agriculture. Basic data mining techniques helped the official of rural housing service to understand and classify the loyal and disloyal customers. Leena Baliga(2000)<sup>7</sup> in her article in Indian Express Newspapers highlights that Citibank, HDFC Bank and ICICI Bank have taken the lead in using data mining along with leading mobile telephony service providers. Rajanish Dass (2006)<sup>8</sup> suggested that data mining techniques can be of immense help to the banks and financial institutions in this arena for better targeting and acquiring new customers, fraud detection in real time, providing segment based products for better targeting the customers, analysis of the customers' purchase patterns over time for better retention and relationship, detection of emerging trends to take proactive approach in a highly competitive market adding a lot more value to existing products and services and launching of new product and service bundles. Madan Lal Bhasin (2006)<sup>9</sup> opined that the leading banks are using data mining tools for customer segmentation and profitability, credit scoring and approval, predicting payment default, marketing, detecting fraudulent transactions, etc. Chase Manhattan Bank in New York, was facing a financial crunch mainly due to constant decrease in the customer base, then the bank used the techniques of data mining to analyze customer profiles and to use them for their benefits and hence chalked out the strategy for the survival and succeed in its attempt. Data mining is also being used by Fleet Bank, Boston, to identify the best candidates for mutual fund offerings. Ebenezer Airiofolo(2010)<sup>10</sup> highlights the need and benefits of data mining in Nigerian Banking Sector. Data mining techniques are helping and will assist more Nigerian banks, telecommunication, insurance and retail marketing to build accurate customer profile based on customer behavior.

## OBJECTIVES OF THE STUDY

- I. To study the scope and feasibility of implementation of data mining techniques.
  - a) To Study the Economic Feasibility of implementation of Data Mining Techniques.
  - b) To Study the Technical Feasibility of implementation of Data Mining Techniques.
  - c) To Study the Operational Feasibility of implementation of Data Mining Techniques.

## RESEARCH METHODOLOGY

The primary data has been collected from employees of public and private sector banks. The questions asked from the respondents and their responses have been shown in Annexure I and apart from the technical questions, the demographic details collected are: Gender, Age, Experience in Banking Sector, Experience of Using Any Software Package, Primary Job Function i.e Marketing, Operations or IT.

The sample has been selected using Stratified Random Sampling Technique. The population has been divided into three categories i.e. employees working at senior, middle and junior levels and then this set of employees have been further divided on the basis of their area of operations that is marketing, operations or information technology. 100 respondents each from public and private sector banks has been collected. For the Purpose of collecting data 100 respondents from Punjab and 100 respondents from Haryana have been considered. It was further studied that among 200 genuine respondents 80 respondents i.e. 40% respondents are from those banks where already data mining concept is in shape and the rest 120 respondents are from the banks which are not using the data mining concept in their banks.

The list of banks included in the survey is:

PUBLIC SECTOR BANKS	PRIVATE SECTOR BANKS
<ul style="list-style-type: none"> <li>• Punjab National Bank</li> <li>• State Bank of Patiala</li> <li>• State Bank of India</li> <li>• Oriental Bank of Commerce</li> </ul>	<ul style="list-style-type: none"> <li>• HDFC Bank</li> <li>• ICICI Bank</li> <li>• IDBI Bank</li> <li>• Axis Bank</li> </ul>

## RESULTS and ANALYSIS

The objective of the study is to investigate the feasibility of implementation of data mining techniques in the banking sector. For this purpose the feasibility has been checked on three major grounds: Economical, Technical and Operational. The questions asked in each sub

section above have been measured on 5 – point Likert Scale i.e. Strongly Agree to Strongly Disagree.

### **Objective I (a): Assessing Economic Feasibility of Implementation of Data Mining techniques in Banks**

The implementation of data mining softwares requires huge investment and the economic feasibility of implementations of data mining techniques in banks is analyzed to check whether the banks are capable enough to spend money on the implementation of these techniques or not. The table I (A)(Annexure I) shows the frequency distribution of response pattern available on various issues regarding economic feasibility. The table shows that the banks are ready to implement the innovate techniques for better functioning and to compete the competition prevailing in the market.

H0: There is no significant difference in approach of both respondents from private and public sector banks regarding economic feasibility of implementation of data mining techniques in banks.

The analysis of the table 1.1 shows that on issues like Bank usually does not hesitate to invest money on the latest hardware ( $p = 0.524$ ), Bank usually does not hesitate to invest money on the latest software's ( $p = 0.162$ ), Bank is ready to invest on IT infrastructure to provide better services ( $p = 0.162$ ), there is no significant difference in approach among the respondents from the private and public sector banks as respondents from each type of bank agree on the fact that their management does not hesitate to invest on latest technologies for the betterment of the organization. Bank will not bother to invest on IT infrastructure to stay ahead of its competitors ( $p = 0.00$ ), on this issues there seems to be difference in approach of the respondents from the both sector of banks. The analysis shows that all respondents from private sector banks agree on all economic issues whereas most of the respondents from public sector banks also agree but some of them are neutral on these issues.

**Table 1.1 Assessing Economic Feasibility (Public Vs Private Sector Banks)**

Banking Sector		D1	D2	D3	D4
Public	Mean	1.73	1.77	1.77	1.44
	Std. Deviation	.660	.678	.678	.585
Private	Mean	1.65	1.65	1.65	1.90
	Std. Deviation	.477	.477	.477	.756
Mann – Whitney Statistics		0.524	0.162	0.162	0.00

In totally, we may conclude that the banks from both the sectors do not hesitate to invest on latest hardware and software infrastructure to increase their work efficiency and to be able to compete with their competitors and to make their stand strong in the market. Although the

management of both types of banks is normally ready to go for investment in the latest infrastructure, but the pace of making the decisions and hence their implementation is fast in private sector banks as compared to public sector banks.

### **Objective I(b): Assessing Technical Feasibility of the implementation of Data Mining Techniques in Banks**

The technical feasibility means whether the banks have or can arrange the required technical infrastructure required for making the data mining techniques functional in the banks or not. This feasibility has been analyzed on questions like whether banks have proper infrastructure to implement latest softwares and hardware and does the bank get the required technical support from the vendors and other agencies to establish the latest software's or not. The table I(B)(Annuxure I) analyzes the frequency distribution of response pattern available for various issues regarding technical feasibility. The analysis of above table show that all respondents agree on the issues that banks have basic infrastructure i.e the required hardware and softwares for the implementation of data mining techniques.

H0: There is no significant difference in approach of both respondents from private and public sector banks regarding technical feasibility of implementation of data mining techniques in banks.

The analysis of the table 1.2 shows that there is significant difference of approach on the issues related to technical feasibility among both private and public sector banks i.e. bank has the basic infrastructure required to use latest softwares ( $p = 0.003$ ), bank has the latest hardware infrastructure to implement the latest softwares ( $p = 0.00$ ), bank gets the required technical support from the vendors and other agencies to establish the latest software's ( $p = 0.00$ ).

**Table 1.2 Assessing Technical Feasibility (Public Vs Private Sector Banks)**

Banking Sector		D5	D6	D7
Public	Mean	2.26	2.26	3.23
	Std. Deviation	.438	.438	.861
Private	Mean	2.14	2.09	2.88
	Std. Deviation	.349	.293	.853
Mann – Whitney Statistics		0.003	0.00	0.00

On technical grounds, the public sector banks are a little bit ahead of the private sector banks and get better support from the vendors. The difference is only due to the difference in the work culture of the private and public sector organizations.

### Objective I(C): Assessing Operational Feasibility of the implementation of Data Mining Techniques in Banks

Apart from economical and technical feasibility there may be some problems at the operational level, the organization may face some hurdles to make these techniques fully operational in the banks. So, the operational feasibility of implementations of data mining techniques in banks has been analyzed on the basis that whether the bank get the required support from the vendors or not, does the bank has sufficient technically skilled man power who can use these techniques efficiently or do they have sufficient technically trained manpower who can train the other employees regarding data mining techniques, does the bank hold training programs for the employees and bank has its own training institute or not.

H0: There is no significant difference in approach of both respondents from private and public sector banks regarding operational feasibility of implementation of data mining techniques in banks.

**Table 1.3 Assessing Operational Feasibility (Public Vs Private Sector Banks)**

Banking Sector		D8	D9	D10	D11	D12
Public	Mean	1.93	3.15	2.00	3.08	2.00
	Std. Deviation	.504	.824	.000	1.000	.000
Private	Mean	1.59	2.86	2.00	2.19	2.00
	Std. Deviation	.492	.787	.000	.586	.000
Mann – Whitney Statistics		0.00	0.001	1.00	0.00	1.00

The analysis of the table 1.3 show that there is significant difference in approach of the respondents from private and public sector banks regarding operational feasibility of the implementation of data mining techniques in the banks on certain issues like bank will get the required training from the vendors/outside agencies for the use of data mining software ( $p = 0.00$ ) as all private bank respondents have agreed on this issue while some public bank respondents are also neutral, bank has the sufficient technically skilled manpower who can train the bank staff regarding the latest software's (like data mining software) ( $p = 0.001$ ) as the average response of the respondents from private sector is 2.86 while that of respondents from public sector banks is 3.15, bank has its own IT training institutes ( $p = 0.00$ ) as the average response of the respondents from private sector bank is 2.19 while that for the public bank is 3.08.

## DISCUSSIONS AND CONCLUSION

The usage of information technology is proving to be a boon for the banking industry. The need of latest techniques and tools is being felt to increase the overall efficiency of the

system. The ever growing competition is forcing the banks to adopt the latest techniques to stay ahead of others. Data mining is the latest technique which can help the banks for providing better customized services to customers, detecting frauds, forecasting future trends and to help the decision makers for making better decisions. In India, the

Banks have realized the need of such techniques and the study conducted here tries to reveal the fact that whether the banks are ready to implement these techniques or not. For this purpose, the economical, technical and operational feasibility is checked and the observations are:

- a) The banks from both the public and private sector are having required infrastructure needed for the implementation of data mining techniques and tools and if any deficiency exists, the management is ready to invest on the latest hardware and software, in order to stay ahead of its competitors. The study concludes that the implementation of these techniques is economically feasible.
- b) As the banks have or can acquire the latest hardware and/or software required for the successful implementation of these techniques and moreover, the banks are able to get the required support from the vendors, so the implementation is technically feasible.
- c) The study reveals the fact that the banks have skilled and technically literate manpower which can use the latest software or can train the other employees of the banks,. Apart from this all the employees of these banks are using one or the other software in daily routine, so it is not very difficult for them to learn this new technology of data mining in order to improve their work efficiency in particular and that of organizational in general. so the implementation is operationally feasible.
- d) The public sector banks have more training institutes as compared to private sector banks.
- e) The banks usually conduct training programs for the employees, so that they can be trained to use the latest softwraes.

## REFERENCES:

1. Danziger, J. N., & Andersen, K. V. (2002). 'The Impacts of Information Technology on Public Administration: An Analysis of Empirical Research From The 'Golden Age' of Transformation'. *International Journal of Public Administration*, 25(5), 591-627.
2. Engler, H., & Essinger, J. (2000). *The future of banking*. UK: Reuters, Pearson Education.



3. Hedelin, L., & Allwood, C. M. (2002). 'IT and strategic decision making'. *Industrial Management + Data Systems*, 102(3/4), 125.
4. H. R. Nemati and C. D. Barko(2003)"Key Factors for Achieving Organizational Data Mining Success". *Industrial Management and Data Systems*, vol. 103, no. 4, pp. 282–292.
5. A. Vasudevan(1999) Report of the committee on technology up-gradation in banking sector" Reserve Bank of India, May 1999, chapter 6.
6. Rob Gerritsen(1999), *Assessing Loan Risks: A Data Mining Case Study*, IT Pro November- December 1999.
7. Report by Leena Baliga in Indian Express Newspaper on 9th Nov, 2000.
8. Rajanish Dass (2006), *Data Mining In Banking And Finance: A Note For Bankers-* Technical note, Note No.: CISG88., April 2006
9. Madan Lal Bhasin (2006), "Data Mining: A Competitive Tool in the Banking and Retail Industries", *The Chartered Accountant* October ,2006.
10. Ebenezer Airiofolo (2010), *A Publication on Data Mining Benefits in Nigerian Banking Sector*, [www.booksie.com](http://www.booksie.com).

## Annexure I

**Table I(A) Response Rate to Assess Economic Feasibility of the Implementation of Data Mining Techniques**

Question Code	Assessing Economic Feasibility	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
D1	Bank usually does not hesitate to invest money on the latest hardware.	36.6%	58.4%	5.0%	--	--
D2	Bank usually does not hesitate to invest money on the latest software's.	35.6%	58.4%	5.9%	--	--
D3	Bank is ready to invest on IT infrastructure to provide better services.	35.6%	58.4%	5.9%	--	--
D4	Bank will not bother to invest on IT infrastructure to stay ahead of its competitors.	45%	39.1%	15.8%	--	--

**Table I(B) Response Rate to Assess Technical Feasibility of the Implementation of Data Mining Techniques**

Question Code	Assessing Technical Feasibility	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
D5	Bank has the basic infrastructure required to use latest software's.	--	80.9%	19.1%	--	--
D6	Bank has the latest hardware infrastructure to implement the latest software's.	--	83.7%	16.3%	--	--
D7	Bank gets the required technical support from the vendors and other agencies to establish the latest software's.	--	36.6%	24.0%	39.4%	--

**Table I(C) Response Rate to Assess Operational Feasibility of the Implementation of Data Mining Techniques**

Question Code	Assessing Operational Feasibility	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
D8	Bank will get the required training from the vendors/outside agencies for the use of data mining software.	30.4%	65.6%	4%	--	--
D9	Bank has the sufficient technically skilled manpower who can train the bank staff regarding the latest software's (like data mining software).	1.2%	30.2%	37.6%	30.9%	--
D10	Bank usually holds training and development programs for its employees to make them more tech savvy.	--	100%	--	--	--
D11	Bank has its own IT training institutes.	--	71.8%	28.2%	--	--
D12	All the staff members are at least that much computer literate that they can use one or the other software banking solutions.	--	100%	--	--	--