



Applications of Data Mining in Banking Sector

Dr. Mohammad Miyan

Associate Professor, Shia P. G. College, University of Lucknow
Sitapur Road, Lucknow (UP) India

Abstract: The data mining (DM) is a great task in the process of knowledge discovery from the various databases. In the corporate sectors, every system has the tough competition with the other system with respect to their value for the business and the financial improvement. Data mining, a dynamic and fast-expanding field, which applies the advanced data analysis techniques, from machine learning, statistics, artificial intelligence or database systems to find out relevant trends, patterns and relations present within the data, information impossible to observe manually.

This paper presents the applications of data mining in the banking sectors. It contains a general overview of data mining, providing a definition of the concept, primary data mining techniques and mentioning the main fields for which the data mining can be applied. It also presents the banking business sector which can benefit from the use of DM tools, along with their use cases i.e., retail and insurance banking sector. Also the commercially available DM tools and their key features are discussed within the paper. With the analysis of DM and the business areas that can successfully apply it, it also presents the main features of the DM solution by using the DM software and data languages i.e., based on survey of users throughout the world that can help to improve the customer experiences and decision making that can be applied for the banking system and the architecture, with its main components for the solution, customer segmentation, banking profitability factor, marketing, risk management and customer relations management and the fraud detection.

Keywords: Architecture; Banking system; Business; Data mining; Data warehouse.

I. INTRODUCTION

In the financial industry throughout the world, the traditional face to face customer contacts are being replaced by the electronic system of contact for reducing the time and cost of the processing an application for the various products and finally improve the financial performance. The use of web, computerization of the financial systems and automated software's has changed the primary concept of the business and the way the successful operation of the business sector. The banking sector has also witnessed an enormous change in the way the banking systems are carried out. Since 1990's, the complete concept of the banking system has been changed in to online transactions, centralized databases and ATM's all over the world that has made banking system technically more strong and customer friendly [1].

The data is one of the valuable resource of any bank but only if it knows how to reveal useful knowledge hidden in the raw data. The data mining permits extracting knowledge from the historical data and also forecasting the outcomes of the future situations. It also helps to maximize the better decisions, increase the value of all the customers, communication and increase the customer satisfaction [1].

So, the Total Branch Computerization i.e., TBC software packages being used at the various branch levels are transaction oriented, as these were designed to keep day by day transactions in mind. The designing of new MIS or restructuring of existing ones could not be possible by replacing the existing TBC packages. The solution appears to be in incorporation of the data warehousing and data mining [1]. Now, due to the huge expansion of the perspective of the data and its multi uses, the systems and individuals are feeling a necessity of a centralized data management and restored system. The centralization of data is required initially for the suitable processing and in turn easy for the user analysis and access [2].

II. DATA MINING

The data mining is the most popular technological tool that converts clumps of data into useful knowledge that can help the data users make informed choices and take suitable actions for the satisfaction. In clear terms, data mining looks for the hidden patterns amongst vast sets of data that can help for understanding, predicting and guiding the behavior in future. The technically, Data Mining is the set of methodologies used in the analysis of data from multi-dimensions and prospects, classifying and grouping the data, finding of prior unknown hidden patterns and precise the identified relations. So that the data mining is in its early stage of development, financial organizations in a broad range of industries including health care, retail sectors, finance, transportation, manufacturing and aerospace, are in the users of data mining techniques and tools to take the huge historical data. By the use and applications of pattern based statistical, mathematical techniques and technologies, to shift via warehoused data observations, DM helps to analysts for knowing significant facts, relationships, trends, patterns, anomalies and irregularities that might otherwise not observed suitably [3], [4].

For the business purpose, data mining is used for finding the patterns and relationships among the data in order to help make better business oriented decisions. The DM can also help to develop some suitable marketing campaigns, spot sales trends and exactly predict the customer loyalty [5].

A. Data Mining Technologies

The analytical techniques used in data mining are frequently well known mathematical algorithms and techniques. What is new is the application of those methods to the business problems made possible by the increased availability of

data, economic storage and processing power. The use of the graphical interfaces has led to tools becoming available that business experts can use easily.

Some important tools used for DM are as follows [6], [7]:

- *Artificial neural networks:* Non linear relative models that learn through training and favorable biological neural networks in the structure.
- *Decision trees:* The tree shaped structures that represent the sets of suitable decisions to make the rules for distinguishing of the dataset.
- *Rule induction:* The withdrawal of useful if then rules from data based on the statistical significance.
- *Genetic algorithms:* The optimization process depends on the logic of alteration, heredity combination and the existing selection.
- *Nearest neighbor:* A method of classifying all the record depends on the previous similar records in the historical database.

The specific uses of data mining are as follows [6], [7], [8]:

- *Customer churn:* Analyzing about the customers those are in the stage of outing and go to other competitor.
- *Trend analysis:* Analyzing the variation between the typical customers in the present month and the previous month.
- *Market segmentation:* Collecting the common criteria of the customers those are interested in the same products of organization.
- *Interactive marketing:* Knowing about every person seeing the website is only interested in seeing or others.
- *Direct marketing:* Analyzing the probability that a person included in the mailing list to get the maximum positive response.
- *Market basket analysis:* Collecting information about the products, which are commonly purchased together; e.g., opening a current A/C and RD A/C.
- *Fraud detection:* Knowing about the transactions within time that are likely to be fraudulent.

DM technology can start new business opportunities as follows [9], [10]:

- *Prediction of behaviors and trends:* the data mining converts the process to find out the result oriented information in the large database. The problems that traditionally required extensive analysis can now be directly answered from the data. A typical case of a predictive problem is targeted marketing. The data mining uses data on past promotional mailings to recognize the targets most likely to maximize return on investment in the future mailings. Other predictive problems include forecasting bankruptcy and other forms of default,

the recognition of segments of the population likely to respond similarly to given events.

- *Discovery of previous unknown patterns:* the data mining tools sweep through databases and identify old hidden patterns. For example the pattern discovery is the analysis of retail sales data to identify apparently unrelated products, which are purchased together. The some pattern discovery problems including detecting the fraudulent credit card transactions and identifying the useless data that can represent the data entry key errors.
- *Using large set of parallel computers:* the companies make an excavation through big amount of data to know patterns about their customers and products. For example, grocery chains have found that when men go to a supermarket to buy shirt, they sometimes walk out with a trouser as well. By taking the information, the products can be arranged accordingly.

III. RELATED RESEARCHES

U. Fayyad et al. (1996), [6] a senior researcher at Microsoft Research, have concluded that the understanding of data mining and model induction at this segment level clarifies the behavior of any DM algorithm and makes it easier for the user to understand its complete contribution and applicability for the knowledge discovery in databases i.e., KDD process; shown by figure 1.

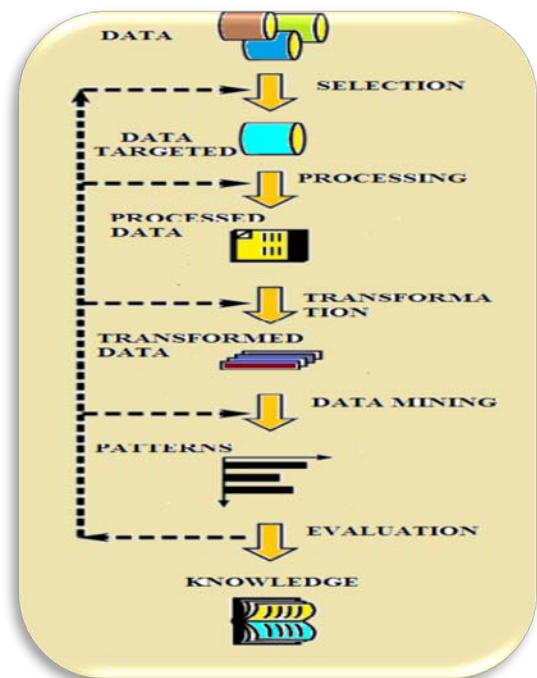


Figure 1 (Various steps of knowledge discovery in databases (KDD) process)

O. Maimon et al., (2010) [8] have discussed the concept of Data Mining and Knowledge Discovery in their hand book. G. S. Linoff et al. [9] (2011) and J. Han et al., [10] (2011) have given the Data Mining Techniques for Marketing, Sales and Customer Relationship Management. V. Bhambri

(2011) [2] has discussed the applications of data mining in the banking sector.

A. B. Devale et al. (2012), [11] have analyzed the applications of data mining in the field of life insurance. I. Lungu et al., (2012) [7] have discussed the Improving Decision Support Systems with Data Mining Techniques and advances in Data Mining Knowledge Discovery and Applications. K. I. Moin et al., (2012) [1] have said that data mining is a tool used to extract some important information from the existing data and enable suitable decision making throughout the banking sectors. R. M. Shafi et al., (2012) [12] have analyzed the DM as a tool for Enhancing Business Process in the banking sector.

V. Jayasree et al., (2013) [13] have given a review literature on the data mining in banking sector. R. Petre, (2013) [5] has said that the architecture proposed for the data mining solution for the business environment would improve the efficiency of the company, by providing valuable decision making knowledge to minimize operating costs and gain competitive advantage.

P. S. Raju et al., (2014) [3] have given the clustering structure and probability density analysis method for detecting the fraud in banks and financial sectors. S. Sharma et al., (2014) [14] have also discussed about the effective uses of DM in the banks.

Sonia et al., (2015) [15] review on Use of Data Mining in Focusing Bank Frauds and Enhancing Business. M. R. Islam et al., (2015) [16] have discussed the DM approach for the prediction of the prospective business sector for lending in the retail banking sector with the use of decision tree. M. P. Thapliyal, (2015) [4] has said that the DM helps to arrange business decisions systematically, detecting the fraudulent transactions, identifying the loyal and better customers, increasing the response of customers, suitable communication and also improve customer's satisfaction.

IV. STATEMENT OF PROBLEM

The efficiency and the performance of the banks are the main sources of the efficacy and efficiency of the countries' financial system. The biggest objective of the banking system reforms has been to increase the profitability, performance and efficiency of the banking sector. The present analysis tries to discuss the problems and reasons related to the marketing, cross-selling, risk management, fraud detection, customer acquisition and retention etc. in the banking system.

V. SCOPE OF THE STUDY

- The present analysis will give a suitable way to the banking system and also the customers about the complete efficiency at which the banks are serving.
- This analysis will also help to know the financial abilities of the banks through DM.
- This study will also spreads light on the different aspects where the banks can work better and how the banks will provide a suitable task to get targeted profit and create a customer friendly good environment to achieve the goal.

VI. DM APPLICATIONS IN THE BANKING SECTOR

The banking industry across the whole world has under gone prodigious changes in the way the business is conducted. By the current implementation, greater acceptance and usage of the electronic banking, the capturing of transactional data has become easier and so the volume of such data has grown. It is not possible manually to analyze this big quantity of raw data and to transform the data into useful knowledge for the banks and financial organization effectively. The huge 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 analyze the patterns and trends, with increased accuracy, which customers will be at a high risk for defaulting on a loan, bank executives can predict, how the customers will react for adjustments in interest rates, which customers will be likely to accept useful offers, and how to make customer relations more profitable. It has among the richest and largest pool of customer information, transactional data, covering customer demographics, credit cards usage pattern, and so on. As banking is in the service industry, the task of maintaining a strong and effective CRM is the critical issue. By using suitable data mining tools, the banks can offer products and services to customers. There are wide areas in which data mining can be used in the banking system, which include customer segmentation and profitability, predicting payment default, credit scoring and approval, detecting fraudulent transactions, marketing, forecasting operations and cash management, ranking investments and optimizing stock portfolios. Therefore, there is a need to build capability to address all the discussed issues and data mining attempts to provide the answer. There are some examples of how the banking system has been utilizing data mining in some areas [1], [2], [4].

A. Marketing Criteria

The bank analysts can analyze the past trends, determine the present demand and forecast the customer behavior of various services and products in order to get more business opportunities and expected behavior patterns. The DM techniques also help to detect the profitable and good customers from the others. The cross selling is the marketing area in which the DM can be used broadly. So, the agency makes it attractive to its customer for buying some other related products or services within the common business framework. The more products and services, that a bank can start for its customers to retain their customers [1], [2], [14].

B. Risk Management Analysis

The DM is also used to minimize the risk in the banking industry. The bank executives want to know whether the customers they are dealing with are suitable, reliable or not. By offering the new customers credit cards, extending the existing customers lines of credit, and approving the loans can be risky decisions for banks as they do not know much about their customers. The DM can also be used to minimize the risk that issue credit cards by considering those customers who are defaulters on their accounts. The customers who have been with the bank for long time periods, remained in the good standing and have higher wages, are more likely to receive a loan than a new customer who has no history with the bank, or who earns low wages. By knowing what the

probability of default is for the customer, the bank is in the good position to reduce the risks [1], [2], [4], [14], [18].

C. Fraud Detection Patterns

Much popular area where the data mining can be used in banking system is in fraud detection. Being able to detect fraudulent actions is an increasing concern for many businesses; and with the help of data mining some fraudulent actions are being detected and reported. The two distinct approaches have been developed by the banks to know about the fraud patterns. For the first approach, a bank bugs the data warehouse of third party and uses the DM programs to detecting the fraud patterns. The bank can then cross reference the patterns with own database for any internal problem. In the other approach, the fraud pattern identification is based on the internal information of the bank by applying a hybrid approach [14], [15], [18].

D. Customer Retention and Accession

The data mining help the banking industry to gain new customers and also help to retain the existing customers. The customer acquisition and retention are very important concerns for an industry, especially the banking system [1], [2], [4].

VII. REMEDIES OF DATA MINING IN BANKS

The remedies of DM in the banks [1], [2] are coming from the need to:

- accommodating of the particular efficiency factor with the prediction accuracy;
- forecasting the multi-variants time series with the sound noise;
- making of the coordinated high resolution forecast;
- having ability to give forecast and the forecasting structure;
- having ability to give benefit from much fine patterns with a small life time;
- incorporate a way of directions as input data for forecasting the structures;
- including the market impact on the concerned persons for market regularities.

A. Time series analysis

A dataset T known as the time series is structured to find out its main components like as Cyclic variation trend $C(T)$, Long term variation trend $L(T)$, Seasonal variation trend $S(T)$ and Irregular movement trends $I(T)$ etc. Let us suppose that T is a time series such as daily closing price of a product P index from moment 0 to current moment M, then the next value of the time series $T(M+n)$ is structured by the formula:

$$T(M+n) = C(T)+L(T)+ S(T)+I(T) \tag{1}$$

The old ARIMA models engaged this area to find out the parameters of the functions by using the formula. The ARIMA models are well known but are difficult to use for the highly non linear statistical processes. The DM methods can be used to make such models to defeat the limitations of ARIMA model. The benefits of quadcore model in comparison with the old black box models such as the neural networks are that components in the formula have an interpretation [8].

VIII. DATA MINING TOOLS

The various data tools are used by the system for data mining through the software. The 16th annual *KDNuggets* software poll was held to get attention from analytics and data mining community, those were using a record number of 93 different tools [17].

The survey analysis of the users of Big Data, Data mining, and Data Science software in the past 12 months for a real project was done amongst 2759 voters. The following table shows the results of the poll in detail % alone is the percent of tool voters used only that tool alone. The result is shown by the table given below [17]:

Table 1 The poll Analysis, Data Mining software used [17] (May 2015)

Legend:	
Red:	% users in 2015
Free/Open Source tools	% users in 2014
Green:	% users in 2013
Commercial tools	
Fuchsia:	
Hadoop/Big Data tools	
R (1293), 3.6% alone	46.9% 38.5% 37.4%
RapidMiner (870), 13.7% alone	31.5% 44.2% 39.2%
SQL (853), 0% alone	30.9% 25.3% na
Python (837), 0% alone	30.3% 19.5% 13.3%
Excel (631), 0% alone	22.9% 25.8% 28.0%
KNIME (553), 6.7% alone	20% 15.0% 5.9%
Hadoop (507), 0% alone	18.4% 12.7% 9.3%
Tableau (341), 0% alone	12.4% 9.1% 6.3%
SAS base (313), 0.6% alone	11.3% 10.9% 10.7%
Spark (311), 0% alone	11.3% 2.6% na
Weka (310), 0% alone	11.2% 17.0% 14.3%
SAS Enterprise Miner (302), 3.6% alone	10.9% 7.2% 5.9%
Microsoft SQL Server (268), 0% alone	9.7% 10.5% 7.0%
MATLAB (243), 0% alone	8.8% 8.4% 9.9%
scikit-learn (229), 0% alone	8.3% na

	na
Unix shell/awk/gawk (221), 0% alone	8.0% 5.8% na
IBM SPSS Statistics (213), 0% alone	7.7% 7.7% 8.7%
IBM SPSS Modeler (197), 7.1% alone	7.1% 5.7% 6.1%
Alteryx (155), 39.4% alone	5.6% 3.1% 10.3%
Pig (150), 0% alone	5.4% 3.5% na
Other programming languages (140), 0% alone	5.1% 3.0% na
Other free analytics/data mining tools (138), 0% alone	5.0% 5.1% 3.4%
Other Hadoop/HDFS-based tools (125), 0% alone	4.5% 3.9% na
TIBCO Spotfire (119), 11.8% alone	4.3% 2.8% 1.4%
Rattle (117), 0.9% alone	4.2% 4.9% 4.5%
QlikView (116), 0% alone	4.2% 3.0% 2.4%
Revolution Analytics (now part of Microsoft) (109), 0% alone	4.0% 9.1% 4.5%
Microsoft Azure ML (102), 1.0% alone	3.7% na na
Microsoft Power BI (98), 0% alone	3.6% na na
MLlib (91), 0% alone	3.3% 1.0% na
JMP (86), 0% alone	3.1% 3.8% 4.1%
SAP (including former KXEN) (82), 26.8% alone	3.0% 6.8% 1.4%
Perl (79), 0% alone	2.9% 3.0% na
Mahout (76), 0% alone	2.8% 2.5% na
Pentaho (74), 0% alone	2.7% 2.6% na
Other paid analytics/data mining/data science software (66), 6.1% alone	2.4% 1.9% 2.4%
Salford	2.3%

SPM/CART/Random Forests/MARS/TreeNet (64), 43.8% alone	3.6% 2.2%
Gnu Octave (64), 0% alone	2.3% 3.9% 2.9%
IBM Watson Analytics (57), 0% alone	2.1% na na
Ayasdi (56), 10.7% alone	2.0% na na
Dataiku (56), 7.1% alone	2.0% na na
Action (56), 7.1% alone	2.0% 0.5% na
H2O (Oxdata) (55), 0% alone	2.0% 0.2% na
Orange (53), 0% alone	1.9% 3.4% 3.6%
Mathematica (52), 0% alone	1.9% 2.3% 2.1%
IBM Cognos (51), 0% alone	1.8% 1.8% 2.4%
Dell (including StatSoft) (47), 19.1% alone	1.7% 1.7% 9.0%
XLSTAT for Excel (42), 0% alone	1.5% 1.2% 0.9%
Stata (36), 2.8% alone	1.3% 1.4% 2.1%
Lexalytics (35), 28.6% alone	1.3% na na
Vowpal Wabbit (35), 0% alone	1.3% na na
C4.5/C5.0/See5 (35), 0% alone	1.3% 1.5% 1.1%
Julia (31), 3.2% alone	1.1% 0.8% na
Splunk/Hunk (30), 0% alone	1.1% 0.7% na
Datameer (26), 0% alone	0.9% 0.8% na
Microstrategy (24), 0% alone	0.9% na na
BigML (23), 0% alone	0.8% 0.9% na
Zementis (22), 31.8% alone	0.8% 0.8% 0.9%
Miner3D (22), 9.1% alone	0.8% 0.9% 1.8%
Oracle Data Miner (22), 0% alone	0.8% 2.2% 1.0%

Amazon Machine Learning (20), 5.0% alone	0.7% na na
F# (18), 0% alone	0.7% 0.5% 0.7%
BayesiaLab (16), 12.5% alone	0.6% 4.1% 1.0%
Dato (former Graphlab) (15), 6.7% alone	0.5% 0.9% na
Clojure (13), 0% alone	0.5% 0.5% na
Alpine Data Labs (13), 0% alone	0.5% 2.7% na
Angoss (11), 18.2% alone	0.4% 0.4% 0.3%
Lavastorm (10), 0% alone	0.4% 0.3% 0.4%
Lisp (10), 0% alone	0.4% 0.3% na
Predixion Software (10), 0% alone	0.4% 3.7% 2.7%
WordStat (9), 0% alone	0.3% 0.2% 0.4%
Megaputer Polyanalyst/Text Analyst (8), 0% alone	0.3% 0.1% 0.1%
WPS: World Programming System (7), 0% alone	0.3% 0.2% na
GoodData (6), 0% alone	0.2% 0.1% na
MetaMind (5), 0% alone	0.2% na na
SiSense (5), 0% alone	0.2% 0.1% na
RapidInsight/Veeva (5), 0% alone	0.2% 0.5% 0.5%
Skytree (3), 0% alone	0.1% na na
Birst (2), 0% alone	0.1% na na
Ontotext (1), 0% alone	0% na na
FICO Model Builder (1), 0% alone	0% 0.2% na

continues to be most popular suite for data mining or data science. It can be noted that a big increase in Hadoop / Big Data tool usage i.e., 29% increase from 17% (2014) and big increase is for Spark whose usage becomes 3-times to previously (2014). The other tools with good growth include Actian, H2O (0xdata), Alteryx and MLib etc. The data mining software usages and used for the year 2015, [17] are shown by the figure 2 and figure 3 respectively.

The participatory survey of the various regions of world was: Europe 38.4%, US & Canada 41.5%, Asian countries 8.2%, Australia & NZ 3.1%, Latin Americans 6.3%, Africa & countries of MidEast 2.5%. In the year 2015, 91% of the total voters used the commercial software and 73% used free software. Around 27% used only commercial software and only 9% used free software. The survey also indicates that 64% i.e., (increased from 49% in 2014) used both free and commercial software [17].

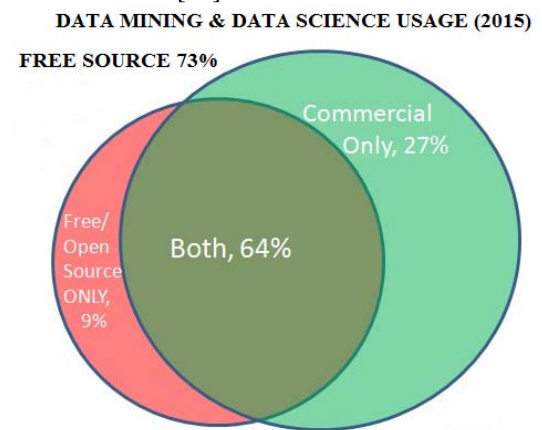


Figure 2

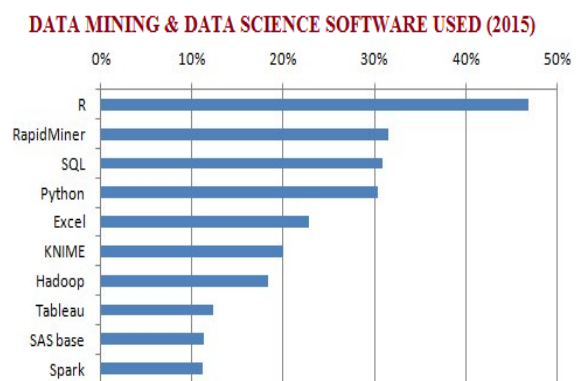


Figure 3

Hadoop /Big Data tool usage increased from 17% (2014); 14% (2013) to 29% among voters. This is due to availability and low cost of many cloud based Big Data tools. The important thing is the increase in Spark usage by 11.3%. Among the programming languages Python is at number one with 30.3% increased usages. The Java is the second most commonly used language with 14.2% usages for analytics/data mining tasks. The tools that showed at least 20% increases in their proportion for two years in the row are Hadoop, Alteryx, Python, KNIME, Qlikview, Tableau, SAS Enterprise Miner and TIBCO Spotfire. Among the tools with at least 20 votes (2014), the big decreasing in 2015, was for these tools, which includes probably a clustering of decline of popularity for free tools like Orange

The R is the most popular overall tool among data miners, although Python usage is growing rapidly. RapidMiner

and lack of a voter drive for some of the commercial tools in the present year [17].

IX. CONCLUSIONS

There are a lot of data mining tools available to give the support for completing the requirements for the banking system. The proposed planning for the DM solution for the banking system can improve the efficiency of the banks and financial sectors, by availing valuable decision making knowledge to minimize the operating costs and gain different advantages from competition view point. The data mining techniques makes the valuable help to the banking system that is related to better growth of the banks by providing them new customers and retaining the old customers by using certain various techniques, how to complete the targets within the certain period of time for the banks and other related system, how to combine the required task in the single bundle known as clustering so that they can be easily performed, providing segmental products for better targeting of customers, fraud detection quickly, analysis of the customers' purchase patterns over time for suitable retention and relationship, detection of emerging trends to take a better approach in a present competitive market adding a lot more value to the existing products and services and initiating the new product and service bundles. The data mining has much wide application domain almost in every field where the data is generated that's why data mining is considered one of the most useful tools in database and information systems and one of the most performing multidisciplinary developments in the field of the banking system.

X. REFERENCES

- [1] K. I. Moin, Q. B. Ahmed, "Use of Data Mining in Banking", International Journal of Engineering Research and Applications (IJERA), Vol. 2, Issue 2, Mar-Apr 2012, pp.738-742.
- [2] V. Bhambri, "Application of Data Mining in Banking Sector" International Journal of Computer Science and Technology, Vol. 2, ISSue 2, June 2011, pp.199-202.
- [3] P. S. Raju, V. R. Bai and G. K. Chaitanya, "Data mining: Techniques for Enhancing Customer Relationship Management in Banking and Retail Industries", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 1, January 2014, pp. 2650-2657.
- [4] M. P. Thapliyal, "Data Mining: A Tool for Banking Industry", International Journal of Emerging Research in Management & Technology, Volume-4, Issue-4, 2015, pp. 84-88.
- [5] R. Petre, "Data Mining Solutions for the Business Environment", Database Systems Journal vol. IV, no. 4/2013, pp. 21-29.
- [6] U. Fayyad, G. P. Shapiro and P. Smyth, "Data Mining to Knowledge Discovery in Databases", AI Magazine, Vol. 17, Issue 3, 1996, pp. 37-54.
- [7] I. Lungu and Adela Bâra, "Improving Decision Support Systems with Data Mining Techniques", Advances in Data Mining Knowledge Discovery and Applications, chapter 18, InTech Publisher, Croatia, 2012, ISBN 978- 953-51-0748-4, pp. 397-418.
- [8] O. Maimon and L. Rokach, "Data Mining and Knowledge Discovery Handbook", Second Edition, Springer Publishing Company, USA, 2010.
- [9] G. S. Linoff and M. J. A. Berry, "Data Mining Techniques: for Marketing, Sales and Customer Relationship Management", Third Edition, Wiley Publishing, USA, 2011.
- [10] J. Han, M. Kamber and J. Pei, "Data Mining: Concepts and Techniques", Third Edition, Morgan Kaufmann Publishing, USA, 2011.
- [11] A. B. Devale and R. V. Kulkarni, "Applications of data mining techniques in life insurance", International Journal of Data Mining & Knowledge Management Process, Vol.2, Issue 4, July 2012, pp. 31-40.
- [12] R. M. Shafi and P. Srinivas, "Data Mining: A Tool for Enhancing Business Process in Banking Sector", International Journal of Scientific & Engineering Research Volume 3, Issue 12, December-2012, pp. 01-06.
- [13] V. Jayasree and R. V. S. Balan, "A Review on Data Mining in Banking Sector", American Journal of Applied Sciences, Vol. 10, Issue 10, 2013, pp. 1160-1165.
- [14] S. Sharma, S. Tarun and P. Sharma, "Effective Use of Data Mining in Banking", International journal of Engineering Sciences & Research Technology, May, 2014, pp. 647-651.
- [15] Sonia, A. Arora, "Review on Use of Data Mining in Focusing Bank Frauds and Enhancing Business", International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 3 Issue VI, June 2015, pp. 175-178.
- [16] M. R. Islam and M. A. Habib, "A Data Mining Approach to Predict Prospective Business Sectors for Leading in Retail Banking Using Decision Tree", International Journal of Data Mining & Knowledge Management Process (IJDMP) Vol.5, No.2, March 2015, pp. 13-22. DOI: 10.5121/ijdkp.2015.5202 13
- [17] KD nuggets News about the software, 2017 KDnuggets. <http://www.kdnuggets.com/polls/2015/analytics-data-mining-data-science-software-used.html>
- [18] M. L. Bhasin, "Data Mining: A Competitive Tool in the Banking and Retail Industries", The Chartered Accountant October 2006. http://www.agricerp.com/mts/data_mining/Banking/banking%20data%20mining.pdf