Mangalayatan Journal of Scientific and Industrial Research

Man. J. Sci. Ind. Res.
Volume 1 (1) (2024), 24105

Article

On Predicting Policy Variables in Banking Sector by Data Mining Techniques

Udita Gupta
Noida Institute of Engineering and Technology
Corresponding author: drudita0108@gmail.com

ABSTRACT

The banking industry has undergone various changes in the way they conduct the business and focus on modern technologies to compete the market. The banking industry has started realizing the importance of creating the knowledge base and its utilization for the benefits of the bank in the area of strategic planning to survive in the competitive market. In the modern era, the technologies are advanced and it facilitates to generate, capture and store data are increased enormously. Data is the most valuable asset, especially in financial industries. The value of this asset can be evaluated only if the organization can extract the valuable knowledge hidden in raw data. The increase in the huge volume of data as a part of day-to-day operations and through other internal and external sources, forces information technology industries to use technologies like data mining to transform knowledge from data. Data mining technology provides the facility to access the right information at the right time from huge volumes of raw data. Banking industries adopt the data mining technologies in various areas especially in customer segmentation and profitability, Predictions on Prices/Values of different investment products, money market business, fraudulent transaction detections, risk predictions, default prediction on pricing. It is a valuable tool which identifies potentially useful information from large amount of data, from which organization can gain a clear advantage over its competitors. This study shows the significance of data mining technologies and its advantages in the banking and financial sectors.

Keywords: *Data Mining, Banking Sector, Fraud Detection, Risk Management, Customer Relationship Management*

1. Introduction

Our present modern era society thrives and evolves on knowledge. Knowledge springs from information gleaned from a good sort of reservoirs of knowledge (databases). Not only does the data itself directly contribute to information and knowledge, but also the trends, patterns and regularities existing within the data files. So it's important to be ready to, in an efficient manner, extract useful information from the info and therefore the associated properties of the info, i.e., patterns and similarities. A new area of research, data extraction or data mining, has evolved to enable the identification of useful information existing within the data reservoirs.

Received: 26 Oct 2024; Accepted: 15 Nov 2024

The newest answer is data mining, which is getting used both to extend revenues and to scale back costs. The potential returns are enormous. Innovative organizations worldwide are already using data processing to locate and appeal to higher-value customers, to reconfigure their product offering to extend sales, and to attenuate losses due to error or fraud.

Data Mining is a process that uses a variety of data analysis tools to discover patterns and relationships in data which will be wont to make valid predictions. Data Mining is that the extraction of hidden predictive information from large databases; it's a strong technology with great potential to assist organizations specialise in the foremost important information in their data warehouses. Data Mining tools predict future trends and behaviors, helps organizations to form proactive knowledge-driven decisions. Data Mining tools can answer the questions that traditionally were too time consuming to resolve. They prepare databases for finding hidden patterns, finding predictive information that experts may miss because it lies outside their expectations.

The first and simplest analytical step in Data Mining is to explain the data- summarize its statistical attributes(such as means and standard deviations), visually review it using charts and graphs, and look for potentially meaningful links among variables(such as values that always occur together).

But data description alone cannot provide an action plan. We must build a predictive model based on patterns determined from known results, then test that model on results outside the original sample. A good model should never be confused with reality, but it is often a useful guide to understanding in banking sector.
Originated from Knowledge Discovery from Databases (KDD), also referred to as Data Mining (DM), it is the nontrivial extraction of implicit, previously unknown and potentially useful information from data in databases. Though, data mining and knowledge discovery in Databases (or KDD) are frequently treated as synonyms, data mining is actually a part of the knowledge discovery process. Data Mining (DM) is that the extraction of new knowledge from large databases. Many techniques are currently utilized in this fast emerging field, including statistical analysis and machine learning based approaches. With the rapid development of the World Wide

Web and therefore the fast increase of unstructured databases, new technologies and applications are continuously coming forth it his field.

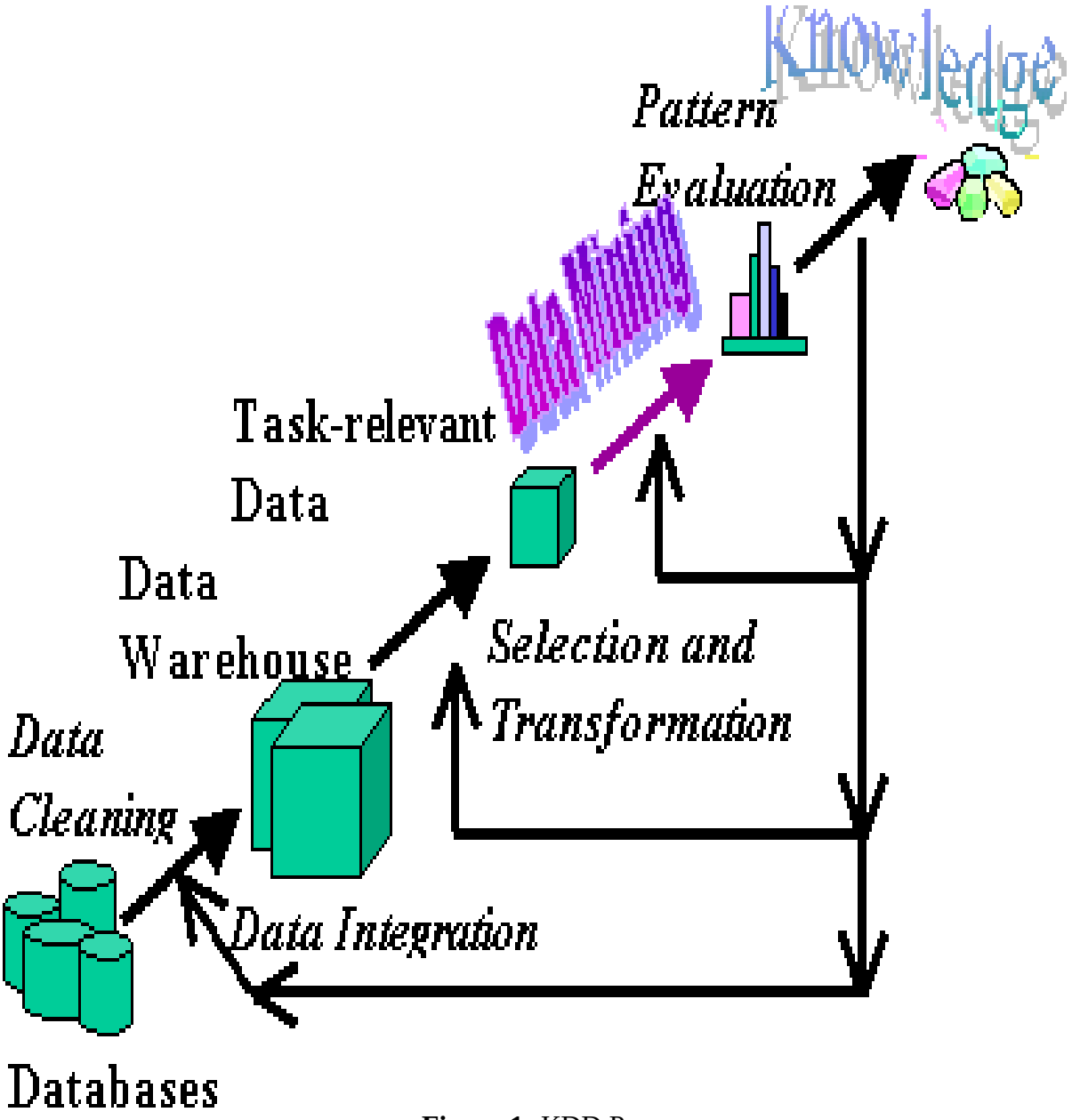

Figure 1: KDD Process

2. The main challenges in Data Mining are:

Data Mining to affect huge amounts of data located at different sites. The amount of data can easily exceed the terabyte limit;

Data Mining is very computationally intensive process involving very large data sets. Usually, it is necessary to partition and distribute the data for multiprocessing to achieve acceptable time and space performance. In many application domain data to be mined either is produced with high rate or they actually come in streams. In those cases, knowledge has to be mined fast and efficiently so as to be usable and updated.

Knowledge discovery in Databases (KDD) is that the process of extracting models and patterns from large databases. The term Data Mining (DM) is usually used as a synonym for the KDD process although properly speaking it is just a step within KDD. DM refers to the method of applying the discovery algorithm to the data. We define the KDD process as: KDD is that the process of model abstraction from large databases and searching for valid, novel, and nontrivial patterns and symptoms within the abstracted model. In this definition, the term process implies that KDD consists of various steps such as: data preparation, search for patterns, knowledge evaluation and refinement. The discovered patterns should be valid with a point of certainty, and novel (at least to the system and preferably to the users). Nontrivial delineates that the discovered patterns should not be obvious within the domain knowledge. They ought to, however, represent a substantial discovery to the user; otherwise the value of the KDD process will not be justified.

Data Mining is additionally stated as essential process where intelligent methods are applied so as to extract the data patterns. Data Mining consists of 5 major elements:

Extract, transform and load transaction data onto the data warehouse system.

Store and manage the data during a multidimensional database system.

Analyze the data by application software.

Present the data during a useful format, like a graph or table.

Data Mining: Convergence of Three Technologies

The convergence of three technologies in Data Mining are:
Increasing Computing Power Statisti-
cal & Learning Algorithm
Improved Data Collection and Management

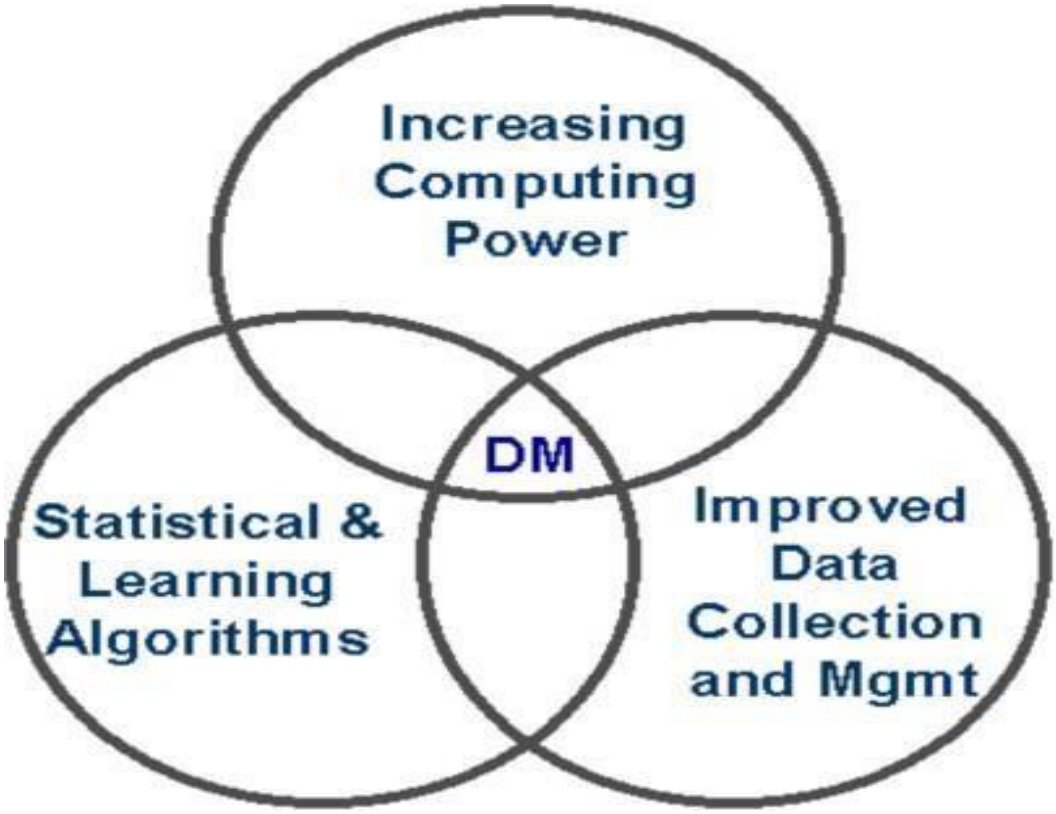

Figure2: Convergence of Three Technologies

3. TYPES OF DATA MINING SYSTEMS:

Data Miningsystems are often categorized consistent with various criteria the classification is as follows:

Classification of Data Mining systems according to the kind of knowledge discovered: This classification supported the type of knowledge discovered or data mining functionalities, like characterization, discrimination, association, classification, clustering, etc. Some systems tend to be comprehensive systems offering several data mining functionalities together.

Classification of Data Mining systems according to mining techniques used: This classification is consistent with the data analysis approach used like Machine Learning, Neural Network, Genetic Algorithms, Statistics, Visualization, Database Oriented or Data Warehouse- Oriented, etc.
This classification also can take under consideration the degree of user interaction involved in data mining process like query-driven systems, interactive exploratory systems, or autonomous systems. A comprehensive system would offer a good sort of data mining Techniques to fit different situations and options, and offer different degrees of user interaction.

4. Data Mining and Banking Sector:

Data Mining has been defined as "the non-trivial extraction of implicit, previously unknown, and potentially useful information from data. It is "the science of extracting useful information from large databases". A.Vasudevan, recommended the utilization of Data Mining techniques, data available at various computer systems are often accessed by a mixture of techniques like classification, clustering, segmentation, association rules, sequencing, decision trees. Trends are often analyzed and predicted with the supply of historical data and therefore the Data warehouse assures that everybody is using an equivalent data at an equivalent level of extraction.

Madan Lal opines that the banks in India and abroad have started using the techniques of Data Mining. Chase Manhattan Bank in New York, Fleet Bank in Boston, ICICI, IDBI, Citi bank, HDFC and PNB in India are using data mining to analyze customer profiles to use them for his or her benefits.

Mudit Saxsena opines that data mining helps to extend sales by targeting the proper customers and to form the proper offers to customers. Banks, who have their ears to the bottom regarding

5. Frauds In Indian Banking Sector

The Reserve Bank of India-RBI maintains data on frauds on the basis of area of operation under which the frauds have been perpetrated. According to such data pertaining, top 10 categories under which frauds have been reported by banks are as follows:

Credit Cards Depos-
its- Saving A/C Inter-
net Banking Housing
Loans
Term Loans
Cheque/ Demand Drafts
Cash Transaction
Cash Credit A/C (Types of Overdraft A/C) Advanc-
es
ATM/ Debit Cards

6. Applications of Data Mining in Banking Sector:

Data Mining can help by contributing in solving business problems by finding patterns, association and correlations which are hidden within the business information stored within the data bases. The banks who have realized the importance of data mining are within the process of reaping huge profits and considerable competitive advantage. According to the regulations given by Federal Reserve Bank of India, the banks need to Provide Off-site Monitoring Surveillance (OSMOS) reports on regular basis in electronic format only and Regulatory requirement of filing of statutory returns like the one under Section 42 of the Reserve Bank of India Act, 1934 for understanding Cash Reserve Ratio (CRR) and Statutory Liquidity Ratio (SLR) obligations in electronic format. According to the Committee formed by Reserve Bank of India Headed by Dr.
A. Vasudevan to travel through the small print of this subject, gave his report on 17th July,1999, the committee highlighted that by the utilization of data mining techniques, data available at various computer systems are often accessed and by a combination of techniques like

classification, clustering, segmentation, association rules, sequencing, decision tree various ALM reports like Statement of Structural Liquidity, Statement of Rate of Interest Sensitivity etc. or accounting reports like record and Profit & Loss Account are often generated instantaneously for any desired period/ date . Trends are often analyzed and predicted with the supply of historical data and therefore the data warehouse assures that everybody is using an equivalent data at an equivalent level of extraction, which eliminates conflicting analytical results and arguments over the source and quality of data used for analysis. In short, data warehouse enables information science to be wiped out a reputable, efficient manner. The Committee recognizes the necessity for data warehouses and data processing both at the individual bank level and at industry level. The implication of adopting such technology during a bank would be as under:

All transactions captured at the branch level would get consolidated at a central location. Such a central location might be called the data warehouse of the concerned bank. For this to happen, one among the wants would be to determine connectivity between the branches on the one hand and therefore the data warehouse platform on the other.

For banks with sizable amount of branches, it is going to not be desirable to consolidate the transaction details at one place only.

7. Application Areas of Data Mining In Banking:

Banking systems contains huge volumes of data. Data's are often both operational and historical. Banks who apply data mining techniques in their decision making hugely benefit and hold a foothold over others who don't. A number of these decisions are within the areas of marketing, risk management and default detection, fraud detection, customer relationship management and concealment detection.

8. Customer Relationship Management:

A CRM system is a sound business strategy for banks to assist create brand value and identify and understand their customers' needs by providing targeted, timely and relevant information which will add value to their customers. CRM systems provide tools which will segment, and deliver the right service, at the right time, by working on dynamic customer information. This enables the power to trace and build strong relationships with profitable customers and identify specific products and services which will benefit customers. Data Mining are often useful altogether the three phases of a customer relationship cycle: Customer Acquisition, Increasing value of the customer and Customer retention. Data Mining technique are often used to create customer profiling to group the likeminded customers in to at least one group and hence they will be dealt accordingly.

In CRM, to get the new customers clustering technique is employed. k- Means may be a distance-based clustering algorithm that partitions the data into a predetermined number of clusters. For example, segment customer profession data into clusters and rank the probability that a private will belong to a given cluster, and provides them banking services they could need.

In CRM, to retain need. The customers, Apriori algorithm are often used to perform market basket analysis by discovering co-occurring items (frequent item sets) within a group. For example, find the things or attributes which comes from the lost customers and specify their association rules. Therefore, the bank can take much care of these customers.

9. Fraud Detection:

Another popular area where data mining are often utilized in the banking system is in fraud detection. Having the ability to detect fraudulent actions is an increasing concern for several businesses; and with the assistance of knowledge mining more fraudulent actions are being detected and reported. Two different approaches are developed by financial institutions to detect

In the first approach, a bank taps the data warehouse of a 3rd party and use data mining programs to spot fraud patterns. The bank can then cross- reference those patterns with its own database for signs of internal trouble. In the second approach, fraud pattern identification is predicated strictly on the bank's own internal information. Most of the banks are employing a hybrid approach.

10. Marketing

As we are already aware of cut throat competition prevailing within the market in most areas, and banking sector is not an exception to it. The marketing and customer care goes hand in hand. Know Your Customer (KYC) is that the buzzword lately. Financial institutions are finding it harder to locate new previously unsolicited buyers, and as a result they're implementing aggressive marketing program to accumulate new customer from their competitors. Moreover the uncertain behavior of the customer is making this task more tedious. A stimulating tool available in marketing and financial organization is analysis of client's data. This allows analysis and calculation of key indicators that help bank to spot factors that affected customer's demand within the past and customer need in the future. Data mining techniques will help in making customer oriented strategies for his or her customers in various categories. The data mining techniques are often used to determine that how customers will react to adjustments in interest rates, which customers are going to be likely to simply accept new product offers, the danger profile of a customer segment for defaulting on loans, etc. The reaction of the customer for the prevailing and new products are often recorded, according to which the longer term strategies are often designed. They will also use the data mining techniques for cross selling. Data mining can improve the response rates within the spam campaigns because the time required to classify the customers are going to be reduced, this successively will increase the revenues, improve the sales department efficiency from the target group.

11. Literature Review:

Mudimigh A. et. al. (2009) In this paper is on extracting knowledge from customer's data through data mining techniques and improving business performance. The authors studied about CRM with Data Mining. Methodology follows the steps: Customer Inquiry, Clusters of customers, Rule Induction Engine, Customer understanding, Action by Organization.

Vivek Bhambari (2011) ATM"s all over the world, which has made banking system technically strong and more customer oriented. Total Branch Computerization (TBC) software packages being used at various branch for daily transactions, Designing the new MIS or restructuring the existing ones would not be possible by just replacing the existing Total Branch Computerization packages. The solution to this problem is to implement the concept of data mining.

Kazi Imran Moin, Dr. Qazi Baseer Ahmed (2012) The wide availability of huge amounts of data and the need for transforming such data into knowledge encourage IT industry to use data mining. The banking industry around the world has undergone a tremendous change in the way business is conducted. The banking industry has started realizing the need of the techniques like data mining which can help them to compete in the market. Leading banks are using Data Mining (DM) tools for customer segmentation and profitability, credit scoring and approval, predicting payment default, marketing, detecting fraudulent transactions, etc. This paper provides an overview of the concept of DM and highlights the applications of data mining to enhance the performance of some of the core business processes in banking industry.

Dr. Chitra K. and Subashini B. (2013) The author includes in this paper is various data mining techniques and its application in banking like fraud prevention and detection, acquiring new customers, customer retention, automatic credit card approval, provide segment based products, marketing and risk management. The paper gives the listing of algorithms of supervised and unsupervised learning. So through mining environment, decision making process becomes very fast.

Raju P. S. et. al. (2014) Data mining and Customer Relationship Management is needed in Banking and Retail Industry. This paper includes various tasks and applications of data mining useful in these industries. The bank and retail industry realizes that data mining is useful process for decision making and gives advantages in competitive environment in the future.

M.Preethi & M.Vijayalakshmi (2017) In the present scenario, banking is an emerging sector where large volumes of electronic data are being maintained. The important task in banking is

handling huge transactional data and making decisions regarding customer retention, fraud detection and prevention, risk and marketing management. But making decisions by manual is time consuming and error prone. To process these data in an effective manner, data mining techniques and methods are pertinent. By using these techniques several interesting patterns and knowledge base can be retrieved. This article analyzes the various data mining techniques and concepts that can be applied to banking sector to enhance its performance.

Veronika Plotnikova, Marlon Dumas & Fredik P. Milani (2019) Data mining and advanced analytics methods and techniques usage in research and in business settings have increased exponentially over the last decade. Development and implementation of complex Big Data and advanced analytics projects requires well-defined methodology and processes. However, it remains unclear for what purposes and how data mining methodologies are used in practice and across different industry domains. This paper addresses the need and provides survey in the field of data mining and advanced data analytics methodologies, focusing on their application in the banking domain. We have also identified various adaptations of data mining methodologies in the banking domain, and noticed that the number of adaptations is steadily growing. The main adaptation scenarios comprise technology-centric aspects (scalability), business-centric aspects (actionability) and human-centric aspects (mitigating discriminatory effects).

Joy Asuni & Kinjalben Koshiya (2022) Today's globalization and intense rivalry have pushed banks to look for ways to differentiate themselves from their competitors. In addition to carrying out practical operational procedures, banks are increasingly using knowledge-based development as a competitive strategy. The capacity to produce, collect, and save data has grown immensely in recent years. This data may include crucial information worthy of enhancement. Bank officials increasingly rely on data mining (DM) techniques due to the widespread availability of massive datasets and the consequent requirement to convert them into knowledge. As a result, the banking sector globally has witnessed significate changes. The banking sector has begun to see the value of data mining in enhancing its ability to compete. Top financial institutions continuously depend on Data Mining techniques for client profiling, credit risk assessment, default prediction, marketing, and fraud detection. This article introduces the reader to the field of data mining. It describes how it has been used in the financial sector to improve the efficiency of specific critical business procedures.

12. Conclusion

This review described that data mining can be a very powerful and helpful tool to extract important and useful information for banking sector from the historical as well as from the current data. Data mining can be used in various fields of banking like Market segmentation by which banks can segment their customers into different groups, direct mail marketing can help the banks to improve their marketing strategy and to increase their business, customer churn to increase the rate of retention of the customers, risk management to reduce the various risks like creditworthiness and fraud detection to reduce the number of fraudulent. Data mining has wide application domain in almost every industry where bulky data is generated and that is why it is consider as one of the most important and promising developments in Information Technology.

Refrences:

1. Albayrak A. S., Yılmaz Ş.K. (2009) "Data mining: Decision tree algorithms and an application on data of IMKB", Süleyman Demirel University The Journal of Faculty of Economics and Administrative Sciences, 14(2), 31-52.
2. Aşan Z.(2007) "Examining the socioeconomic characteristics of customers using credit cards, with clustering analysis", Dumlupınar University The Journal of Social Sciences, 17 (4), 256-267.
3. Atbaş A. C.(2008) "A study on determining the cluster number in clustering analysis, Master Thesis, Ankara University", Graduate School of Natural Sciences
4. B.Muneendra Nayak, Naveen kumar, R.Mahammad Shafi.(2013). "A Focus On Different Frauds And Using Data Mining To Enhance Business Process In Banking Sector" International Journal of Engineering Sciences Research 04, Special Issue 01.
5. Bilen H.(2009), Data mining application for personnel selection and performance evaluation in banking sector, Master Thesis, Gazi University, Graduate School of Natural and Applied Sciences
6. Bharati M. Ramageri. (2009) "Data Mining Techniques And Applications", Indian Journal of Computer Science and Engineering 1 (4), 5321-5330.
7. Burez, J. and van den Poel, D. (2009). Handling Class Imbalance in Customer Churn Prediction. Expert Systems with Applications, 36(3):4626-4636
8. Chien C.-F., Chen L.F.(2008), "Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry", Expert Systems with Applications, 34(2), 280-290.

9. Ching W. K., Pong M. K..(2002), “Advances in data mining and modeling”, 1st ed., World Scientific Journal , 24(1), 270-290.
10. Chopoorian, J.A., Witherell, R.Khalil, O.E.M and Ahmed (2003) “Mind your business by mining your data”. SAM Advanced Management Journal 34(2), 280-290.
11. Doğan B.(2008), “Clustering analysis as a tool under the supervision of banks: An application for Turkish banking sector”, PhD Thesis, Kadir Has University, Graduate School of Social Sciences.