

Data Mining In Banking Sector

M.Preethi

Assistant professor, Department of Computer Science
Madurai Sivakasi Nadars Pioneer Meenakshi Women's College, Poovanthi
Email:preethi.dec30@gmail.com

M.Vijayalakshmi

Assistant professor, Department of Computer Science
Nadar Mahajana Sangam Sermathai Vasam College for Women, Madurai
Email : mviji2803@gmail.com

ABSTRACT

Data Mining plays an imperative role in many organizations. It is helpful in scrutinizing the collected data and delivering it into an understandable pattern. 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.

Keywords: Banking, CRM, Data Mining, Fraud detection, Knowledge discovery.

1. Introduction:

The banking industry is highly competitive. It is sensitive to political and economic conditions in their domestic countries and all over the world. Because of a lot of risks, a key strategy for many banks is to improve their performance by reducing costs and increasing revenues. On the other hand, as the customer's transactions increased, data stored in computer system are relatively high. In banking sector, data act as an asset of an organization since valuable knowledge and interesting patterns are hidden in that. In conservative settings, decision making will be done by manual procedures. Users may also use many tools for analyzing data at critical decisions. Manual analysis may not be accurate since large volume of data can be analyzed only in a limited manner.

In Figure 1 shows conservative decision making process. To take an effective decision in various areas like Marketing, Risk management, Fraud detection and customer relationship management, data mining techniques are appropriate. Figure 2 depicts the decision making with data mining.

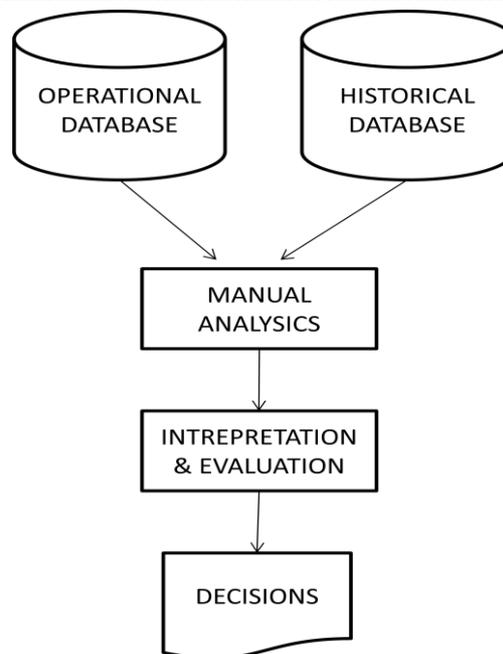


Figure 1 Conservative Decision Making

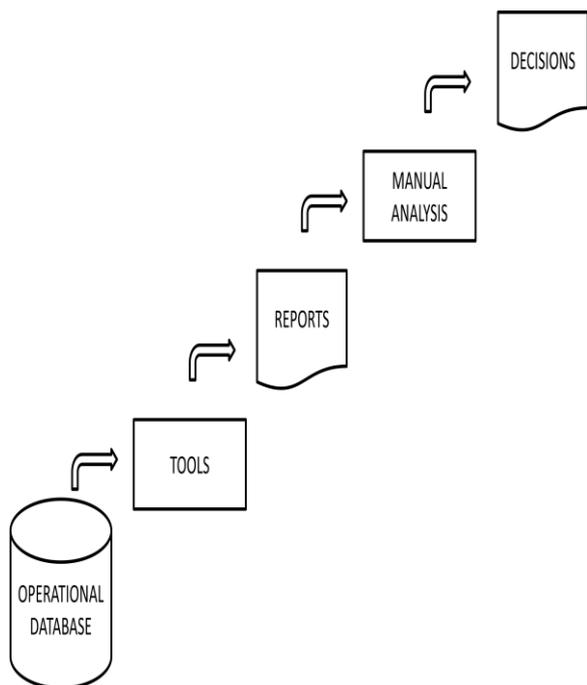


Figure 2: Decision making with data mining.

2. Data Mining and Knowledge Discovery:

Data mining sometimes called data or knowledge discovery is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.

There is a huge amount of data available in the Information Industry. This data is of no use until it is converted into useful information. It is necessary to analyze this huge amount of data and extract useful information from it.

Data mining also involves other processes such as Data Cleaning, Data Integration, Data Transformation, Data Mining, Pattern Evaluation and Data Presentation. Once all these processes are over, we would be able to use this information in many applications such as Fraud Detection, Market Analysis, Production Control, Science Exploration, etc.

2.1 Data Selection

This is the first step in data mining process where data is collected from various databases and warehouse.

2.2 Data Preprocessing

It involves Data Cleaning and Data Integration.

2.3 Data Cleaning

In this stage, inconsistent and irrelevant data are removed.

2.4 Data Integration

In this, various heterogeneous data sources are combined.

2.5 Data Transformation & Data Reduction

Data are transformed by performing aggregation operations. Data Reduction is removing redundant and highly correlated data items.

2.6 Data Mining

In this step, data mining techniques are applied in order to extract data patterns.

2.7. Pattern Evaluation

In this stage, the patterns are evaluated for their relevance implementation.

2.8. Knowledge Presentation

Here visualization and knowledge representation techniques are used to present mined knowledge to the user.

3. Data Mining Techniques:

3.1 Association:

Association and correlation is usually to find frequently used data items in the large data sets. It is the technique of finding patterns where one event is connected to another event. This type of findings help businesses to make certain decisions regarding pricing, selling and to design the strategies for marketing, such as catalogue design, cross marketing and customer shopping behavior analysis [2].

3.2 Clustering:

Clustering can be said as identification of similar classes of objects. This is the technique of combining the transactions with similar behavior into one group, or the customers with same set of queries or transactions into one group. Classification approach can also be used as effective mean of distinguishing groups. So clustering can be used as preprocessing approach for attribute subset selection and classification [1].

3.3 Forecasting:

Regression technique can be adapted for prediction. Regression analysis can be used to model the relationship between one or more independent variables and dependent variables. In data mining independent variables are attributes already known and response variables are what we want to predict [2]. Unfortunately, many real-world problems are not simply prediction. For instance, sales volumes, stock prices, and product failure rates are all very difficult to predict because they may depend on complex interactions of multiple predictor variables [1,8]. Therefore, more complex techniques (e.g., logistic regression, decision trees, or neural nets) may be necessary to forecast future values. This technique of data mining will help in discovering patterns from which one can make reasonable predictions

3.4 Classification:

Classification is the most commonly applied data mining technique, which employs a set of pre-classified examples to develop a model that can classify the population of records at large. Fraud detection and credit risk applications are particularly well suited to this type of analysis. This approach frequently employs decision tree or neural network-based classification algorithms. The data classification process involves learning and classification. In Learning the training data are analyzed by classification algorithm. In classification test data are used to estimate the accuracy of the classification rules [2,3].

4. Application of Data Mining Techniques in Banking:

4.1 Customer Relationship Management:

A CRM system is a sound business strategy for banks to help create brand value and identify and understand their customers' needs by providing targeted, timely and relevant information that can add value to their customers. CRM systems provide tools that can segment, and deliver the right service, at the right time, by acting on dynamic customer information. This allows the ability to track and build strong relationships with profitable customers and identify specific products and services that can benefit customers. Data mining can be useful in all the three phases of a customer relationship cycle: Customer Acquisition, Increasing value of the customer and Customer retention [4]. Data mining technique can be used to create customer profiling to group the likeminded customers in to one group and hence they can be dealt accordingly [2].

✓ In CRM, to discover the new customers clustering technique is used. k-Means is 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 an individual will belong to a given cluster, and give them banking services they might need.

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

4.2 Fraud Detection:

Another popular area where data mining can be used in the banking industry is in fraud detection. Being able to detect fraudulent actions is an increasing concern for many businesses; and with the help of data mining more fraudulent actions are being detected and reported. Two different approaches have been developed by financial institutions to detect fraud patterns. In the first approach, a bank taps the data warehouse of a third party and use data mining programs to identify 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 based strictly on the bank's own internal information. Most of the banks are using a hybrid approach [5].

5. Conclusion:

Data mining is an efficient tool to extract knowledge from existing data. In Banking, data mining plays a vital role in handling transaction data and customer profile. From that, using data mining techniques a user can make a effective decision. Two major areas of banking application are Customer relationship management and Fraud detection. Here the main techniques used are clustering and Association methods. In this paper, based on these methods various types of algorithms are discussed. Finally we conclude that Bank will obtain a massive profit if they implement data mining in their process of data and decisions.

References:

- [1] Hillol Kargupta, Anupam Joshi, Krishnamoorthy Siva Kumar, Yelena Yesha, "Data Mining: Next Generation Challenges and Future Directions", Publishers: Prentice-Hall of India, Private Limited, 2005.
- [2] S. S. Kaptan, N S Chobey, "Indian Banking in Electronic Era", Sarup and Sons, Edition 2002.

[3] S.S.Kaptan, "*New Concepts in Banking*", Sarup and Sons, Edition, 2002

[4] Rajanish Dass, "*Data Mining in Banking and Finance: A Note for Bankers*", Indian Institute of Management Ahmadabad.

[5] Dr. Madan Lal Bhasin, "*Data Mining: A Competitive Tool in the Banking and Retail Industries*", The Chartered Accountant October 2006

[6] R. Groth, *Data mining: building competitive advantages*, Prentice-Hall, 2001.

[7] William W. Cooper, Lawrence M. Seiford, Kaoru Tone, *Data envelopment analysis: a comprehensive text with models, applications, references and dea-solver software*, Kluwer Academic Pub, 1999.