Credit Risk Management of Customers at Iranian Banks by Using data-mining

Rosa Khezri
Student of eMBA, Islamic Azad University (rosa_khezri@yahoo.com)

Dr. Mansooreh Aligholi
Faculty Member of Azad University, (m.aligholi@yahoo.com)

Abstract

Today, governmental supports provide the survival of most Iranian banks and continuation of their activities. Lack of suitable and expertise models within the credit banking network has led to remarkable volumes of the unpaid or delayed granted facilities of the banks. Risk reduction and control are as factors affecting credit granting process improvement and playing significant roles in facility granting, profitability, and viability of the banks and financial institutes. This research aimed at the identification of an appropriate model to classify bank customers and provide a suitable ground to predict the behaviours of prospective customers. To have a more efficient credit risk management within intelligent systems, such as decision support systems, a model was presented by classifying customers based on 2 different perspectives, i.e. Rial and foreign exchange behaviours of customers and banks' shares of facilities and using data mining tools on the information of 319 customers of different operating banks within the period of 2002-2015.

Keywords: Credit Risk, data-mining, Iranian Banks.
Introduction:

One of the important factors in the health of a society's economy is regular and accurate function of cash flow cycle between banks and credit customers since the mentioned cycle will be diminished and practically led to an inappropriate return in case of the retention of resources by the customers. The proper relationship between the financial and productive systems of any country is considered as the major factors of the economic growth and development. According to the economic structure and for reasons such as lack of developed capital markets and other non-bank and contractual networks in Iran, the banking network has to undergo the financing of the real economic sectors. Unfortunately, the banks have not been very successful in achieving this important mission. Now, the continuity and survival of the majority of the country's banks are dependent on governmental supports. A considerable amount of the banks' granting facilities balanced or pended indicate a lack of appropriate and expertise model in the credit systems of Iranian banking network. Anyway, lack of currency system and credit granting has more or less affected all the economic actors and caused economy to encounter serious problems and complexities as heavy surges layer by layer. In a healthy economy, all the economic actors benefit from the proper function of financial systems proportionate to their shares, the condition which will not be established except with the involvement of the same economic actors. Although risk managements of banks and financial institutions have grown late, it seems that they are passing through a transitional period between ordinary banking and novel and scientific banking. Entry into the competitive business environment with new conditions is an inevitable phenomenon, while management in this different environment and condition demands for a more distinct knowledge and experience than common knowledge of management and organization.

From among the risks that the banks face, valuation and measurement of credit risk are of special positions (Zekavat, Morteza, 2008). Reducing and controlling risk have been raised as one of the effective factors on the improvement of the process of granting credits and thus the performance of banks, while having essential roles in the continuity of providing facilities and profitability and survival of banks and financial institutions (Fallah Shams, Mirfeiz&Rashno, Mehdi, 2007). One of the suitable techniques in this field is using decision-making models. Application of these techniques and more practical and consequently mechanized and systematized planning of them can help many banks to manage the crises and risks, including credit risks (KhaliliAraqi, Maryam, 2005). Elimination of credit risk is theoretically and practically impossible; however, a fundamental action can be taken to manage this type of risk by adopting appropriate strategies. Leaders of economic enterprises and their senior advisors need to update management knowledge and business with this approach. Using data-mining tools, customer classification was achieved based on credit facilities in this research. The research method was based on gathering and refining customers' data in accordance with a standardized data-mining process, assessing the different aspects of customers' classes according to bank experts' views and their verbal terms for these variables, and then modelling the final data using different techniques. Thus, managers, especially those of the banks, should learn how to use modern tools and technology to support decision-making.
Research literature:

Data-mining:

The information known as knowledge discovery is mostly derived from data-mining. Data-mining is related to the discovery patterns and the rules or models based on a set or sets of data. The results (patterns, rules, and models) will be used to predict future outcomes. Only those results that happen outside of the prediction limits and patterns require auditors' examinations (Delen and AlHawamdeh, 2009).

Data-mining is defined as the process of obtaining the unknown but understandable data, which are reliable, efficient to be used in a commercial decision-making.

The next decade is going to confront a technological revolution caused by data-mining novel knowledge as one of the 10 developing sciences. On the other hand, there is a more need for companies to understand the hidden part of in their data due to the increasing competitions and thus more investments on the data-mining projects have been conducted in recent years.

Decision-making trees:

The dependent and independent variables include previously known features and future predictions, respectively, when extracting data. However, many of the real issues cannot be easily predicted and thus prediction of future values may depend on more complex separations, one of which is to make a decision tree to classify samples growing from roots and arriving eventually at leaf nodes. A question about the input example is raised by each interior or non-leaf node and classes are represented by the leaf nodes (Ghazanfari, 2013).

To analyze data based on features, an utmost advantage can be taken by C5.0, serving as an algorithm to build a decision tree or a set of rules. According to a feature, each sub-section is divided into a new section until no new ones are divided. However, only one classification target can be predicted by this model (Shahrabi&Zare, 2013).

Risk management:

In the discussion of risk management, a criticism is raised when referring to obscure and inconsistent application of the term 'risk'. One of the most prestigious global forums on banking is Basel Committee, which has no legal power though a majority of its member are implicitly obliged to implement its recommendations (Marrison, 2002).

Reliability of customers and having complete information about them are important for bank managers and decision-makers; otherwise, their loans are accompanied with risky decisions.
Credit risk:

There are several categories of risks. The banking system can be affected by 8 different categories as follows: security risk, legal and ethical risk, traditional risk (including credit risk), reputational risk, operational risk, strategic risk, cross-border risk, market risk, money laundering risk, interest rate risk, and liquidity risk (Zarei, 2011).

Determination of the recipient’s ability to repay principal and interest and the credit rating is important when granting facilities. A credit is defined as the probability of failure to return the principal and interest of loans. Nevertheless, several techniques are employed to measure, manage, and reduce it (Zekavat&Morteza, 2008). An essential condition to the long-term success of any banks is provided through credit risk management as part of a comprehensive management.

Customer credit classification:

Using an applicant’s present and past information, banks and credit institutions can assess and score his or her probability of failure to repay a loan through a system called customer credit classification.

High control and accuracy and low cost are associated with the commercial banks’ credit assessment model based on C5.0 algorithm compared to other patterns (Pang Su-lin, 2009).

It is clear that the central bank's major approach is to encourage credit institutions to create a classification and internal ranking system complied with international standards.

Credit risk management:

The overall strategy of a credit institution together with the credit risk strategies and policies and management practices provide a background for the supervision of the Board of Directors and senior management. An institutional credit strategy mainly aims at determining its risk-taking amount.

To recognize, accept, measure, monitor, and control a credit risk, systems and approaches are required, by and through which examination of a counterparty’s good reputation and credibility in the market and his or her ability to fulfil the commitments will be possible (Khalili Eraqi, 2005).

The purpose of this study was to present a good model of customers' classification to contribute to the credit risk management in Iranian banks using data mining techniques and algorithms. But, is it plausible to classify the Iranian bank customers based on credits and bank's share by using data mining and how is it performed? How important are the effective parameters on bank customers' classification?
Conceptual model:

Based on the studies of Tian-Shyug Leea and his colleagues (2004) and that of Lin and Zhang (2009) carried out in the Institute of Financial Engineering, School of Management, University of Jinan Guangzhou in 2009 and what was said above, a conceptual model was adopted.

Using classification, multivariate adaptive regression, and regression tree, Tian-Shyug Leea and his colleagues (2004) studied customer’s credit information and stressed on the importance of validation between important competing financial institutions. On this basis, more banks are employing credit scoring models to find better strategies daily. In an attempt to extract customer’s credit criteria for an appropriate classification and introduce a validation model of customer’s status, the mentioned researchers made a review of the Client’s File 8000 of the local bank in Taipei, Taiwan, which had received a credit card. They stated that the risk degree of paying loans to customers can be determined via a credit classification as a statistical analysis system. Thus, the rates of risks and errors influenced by various characters and attributes of applicants can be specified for banks to lend more confident facilities (Tian-ShyugLeea, 2004).

As for this study, the 2 following variables were selected after reviewing the previous research and their objectives:

1. Customer’s behaviour data, including application type, contract type, sum of the past due principal balance, sum of the past due interest balance, sum of the delayed principal balance, sum of the delayed interest balance, sum of the delayed past due balance, and returned check.

2. Bank’s share consisting of its accessible credit share percentage divided into the 3 categories of low risk (0% to 30%), medium risk (30% to 60%), and high risk (65 to 100%).

Definition of variables:

Bank’s share, customer type, customer’s number, request type, contract type, sum of the past due principal balance, sum of the past due interest balance, sum of the past due balance, sum of the delayed principal balance, sum of the delayed interest balance, sum of the delayed balance, and returned check.

Using CRISP-DM process model and comparing the resulting decision trees of C5 algorithm, a framework was proposed for bank customers’ credit classification and credit risk management to discover the existing hidden patterns in the bank’s available transactional database based on a set of rules in this research.
Methodology:

This is an "applied" study based on objective and a "descriptive-analytic" research based on data collection method. The statistical community of this study included randomly selected clients of all the agent banks' branches from 2009 to the first quarter of 2015. For the classification of Iranian bank customers, a data mining approach was employed. 319 statistical samples were selected using Cochran formula. This research was mainly based on the knowledge discovery of the study agent banks’ databases through the global standard process.

By applying a bottom-up approach, it was tried to discover the unknown data and make some laws for them. In this study, after refining the data obtained from the desired bank customers' databases, the effective factors on the customers' classification were discovered by interviewing with the experts and collecting scientific evidence. Regarding the sample customers' good and bad payments, a classification label was considered for all of them. Finally, based on the defined classes, their rules and patterns were identified and a framework was presented for predicting the applicants’ credits. Also, to determine the client type based on good and bad payments, a model of hierarchical analysis was utilized.

Analysis method:

In this research, CRISP-DM standard model that provides an overview of the life cycle of a data mining project consisting of the 6 stages of business understanding, data understanding, data preparation, result evaluation, modelling, and model application was employed.
Clementine 12 software was used to apply the methods and CRT, CHAID, and C5.0 algorithms with the highest accuracies and least errors were separately classified to model and extract the knowledge according to their training data.

Results:

To classify the customers comprehensively, their behaviours in terms of monetary and foreign exchange were ranked and bank’s share of risk was reviewed.

Maximum and minimum frequencies of good- and bad-paid customers were resulted by the descriptive findings of their credit-based behaviours.

- The good-paid customers were found to have high (15%), medium (6%), and low (79%) risks based on the bank's share.
- The medium customers were with 2% of high and 98% of low risks based on customer type and bank's share.
- All of the bad-paid customers were with high risks.
- The clients were found with high (62%), medium (approximately 7%), and low (31%) risks according to the customer type, bank's share, and risk.

C5.0, CRT, and CHAID models were selected to find the optimal one for the present study. To this goal, the following table of data analysis was prepared. Finally, the classification results obtained by the models were compared with the help of different analysis techniques and tools.

The results of data analysis shown that C5.0 model has the least errors with a 100% correct percentage and is thus suitable for determining customers' new classes and types.

The 3 good, medium, and bad bank customers were classified based on application type, contract type, sum of the past due principal balance, sum of the past due interest balance, sum of the past due balance, sum of the delayed principal balance, sum of the delayed interest balance, sum of delayed balance, sum of delayed past due balance, and returned check.

C5.0 model considered important the sum of past due principal balance when selecting customer type. It divided it into the 2 groups of ≤ 0 and > 0 for almost 99% and 100% of the medium good-paid customers, respectively.

The group of the sum of the past due principal balance is divided into the sub-group of sum of the delayed principal balance. In this sub-group, 100% of the bad- and good-paid customers were with and without delayed principals, respectively.

Furthermore, 72.1%, 26.95%, and 0.94% of customers were considered to be good, medium, and bad, respectively.
Another factor (bank’s share) was added to the test for choosing a suitable model based on customers’ monetary and foreign exchange behaviours and determine their classes and types. By this, 15.2%, 5.6%, and 79.1% of the good customers were found to be with high, medium, and low risks, respectively. Also, the medium clients were with high (97.67%) and low (2.3%) risks. All the bad customers had high risks. In fact, a few number of good-paid customers were observed to have low banking risks.

After the addition of bank’s share, a correct and suitable relationship was found between all the bank customers based on C5.0 model data analysis, while the rate of error was zero (wrong = 0%, correct = 100%).

Based on the above, the factor of bank's share was regarded to be effective by the branch of sum of the past due principal balance. A bank’s share of 80 belonged to good customers, while a share of over 80 was needed for the sum of delayed principal balance to be effective for 100% of the bad- and good-paid customers, who were with and without delays, respectively.

Assessment of customer’s behaviour type and bank's share in Iran will result in a proper model. Accordingly, high banking risk and low share were obtained for most people in the good-paid group and within the bank’s share group, low risk and high banking risk were achieved for most medium customers. However, within the bad-paid group of customers, no significant differences were discovered.

Following the addition of bank risk to the previous data in C5.0 model, a total of 81.33% of the bank customers in the training and producing groups were found to have a suitable and correct relationship with an error rate of 18.67% (Wrong = 18.67%, correct = 81.33%).

Therefore, CHAID model was selected as a suitable model with higher percentage of accuracy and fewer errors than C5.0 model after performing further tests.

As shown in Figure 3, the model will contribute to the banks' credit risk management of foreign currency determine their micro- and macro-credit strategies.
Conclusion:

Based on businesses, severe changes are occurring to the world's banking industry. The importance of time and speed of access to information has doubled due to the competitions in the scientific, social, economic, political, and military fields. Today, attracting deposits and granting loans have made the core expertise of commercial banks. Banks are exposed to credit risks by granting loans. There is a necessity for credit risk management to attract new customers and keep old customers in the banks. Thus, application of new technological solutions, such as customer classification has emerged to identify target customers and discover knowledge from database using data mining and credit risk management. Data mining tools are used by the leading banks to classify credits and determine the rates of credit risk. As a result, improved management,
credit risk control and reduction, customer selection, granting of banking facilities, and enhancement of productivity level are achieved.

The answers to the research questions are as follows:

The customers can be classified based on credits and a more effective credit risk management is obtained by using the data mining model of bank's clients’ information. The sum of past due principal balance and consequently, the sum of delayed principal balance are affected by the customer type. Customer type is under the influence of the sum of past due principal balance affected by bank’s share, which is influenced by the sum of delayed principal balance. It is concluded that customer’s behaviour depends on the intra-bank factors and affects risk management in the banks of Iran.

The practical results of the research are as follows:

- To decide on the new granting credits and reduce damage to the Iranian banking system, customers' classification should be established based on rial and foreign currency rates derived from the central bank's comprehensive and integrated database of customers’ information available at different agent banks. The mentioned database is obtained through experiences and data mining of customers’ information. As a result, the impacts of personal views, decisions, and violations are reduced and the numbers and amounts of past due and delayed loans are minimized. Nevertheless, customers’ satisfactions of the banks must be taken into account.

- It is inevitable to keep pace with the new knowledge and strengthen the banks’ ITs to effectively manage credit risks.

- The Iranian agent banks’ individual decision-making on granting credits will be minimized and their interests will be improved if they try to seek for new techniques and establish a system of proposals to develop proper legislation. Consequently, they can manage their own credit risks more effectively. Finally, it is highly suitable to establish a system of recommendations to take benefit from the experts’ views of different banks, while regarding each bank's specialized aspects.
References


Fallah Shams, Mirfeiz and râşnov, Mehdi. Credit risk management in banks and credit and financial institutions (concepts and models). Faculty of Economic Sciences, 2007.


Tian-ShyugLeea, Chih-Chou Chiub, Yu-Chao Chouc, Chi-JieLud(2004), Mining the customer credit using classification and regression tree and multivariate adaptive regression splines, pp1113 – 1130

Pang Su-lin, GONG Ji-zhang (2009), C5.0 Classification Algorithm and Application on Individual Credit Evaluation of Banks, 29(12): 94–104