Journal of Money and Economy
Vol. 14, No. 2, Spring 2019
pp. 177-204

Evaluating the Application of a Financial Early Warning System in the Iranian Banking System

Seyed Babak Ebrahimi[*] **Pouria Rezaei**[†]
Pejman Azin[‡]

Received: 21 Jan 2019 Approved: 29 Feb 2020

One of the significant problems of banks and investors in Iran is the lack of precise awareness about the financial performance of each bank and the roadmap for improving the conditions. Besides, the undesirable status of the financial performance of banks becomes evident only when the improvement of conditions is complicated. In this paper, a data mining-based early warning system (EWS) model has been presented to capture the financial performance of banks. To design this model, the CHAID decision tree has been used. Using this model, the banks have been classified as poor, medium, and good regarding financial performance, and the roadmap to achieving the desirable status has been determined. For this purpose, 13 Iranian banks have been investigated within the years 2003-2017. Eventually, the results obtained from the decision tree have been compared with the findings achieved from the CAMELS model. Based on the designed decision tree, 8 profiles have been extracted; 2 representing good, 3 medium, and 3 poor financial performance. Based on these profiles, according to the latest reports published by the studied banks, eight banks have a mediocre financial performance while five banks suffer poor financial performance. According to these profiles, four variables of the asset to shareholders' equity ratio, the shareholders' equity to loans ratio, the long-term debt to equity ratio, and liquidity coverage ratio were identified as the most relevant variables associated with the financial performance of banks.

Keywords: Decision Tree, CHAID Algorithm, Data Analysis, Early Warning System, Financial Risk, CAMELS Model.
JEL Classification: G20; C80

1 Introduction

Financial stability is a relatively new term that has emerged following the extensive and progressive incidence of financial crises in Latin America, southeastern Asia, Russia, and Turkey over the past two decades by the

[*] Department of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran; b_ebrahimi@kntu.ac.ir
[†] Faculty of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran; Ayandeh Bank, Tehran; pooria_rezai12@yahoo.com (Corresponding Author)
[‡] Faculty of Industrial Engineering, Sharif University of Technology, Tehran, Iran; Ayandeh Bank, Tehran; pejmanazin@gmail.com

policymaking and economic literature communities. Adoption of preventive measures to tackle systematic risks in the financial sector and consolidation in financial institutes to reduce the cost of financial crises constitute the main elements of the policy package for financial stability. The financial crisis refers to a situation where a considerable percentage of the value of some assets is lost unexpectedly. Historical evidence suggests that many financial crises are due to the crisis in the banking system, which has eventually led to the economic recession and unemployment crisis (Sotudehnia and Abedi, 2013).

One of the indicators used for investigating the extent of the financial stability of a country is the economic risk index. Table 1 compares the economic risk index of Iran with that of the neighboring countries. This index is presented annually by the PRS company in the international guideline report for the risk of countries to investigate the level of risk in 100 countries of the world. Naturally, the more powerful the economy of the country, the lower their economic risk will be. When calculating this index, the following components have been considered:

- Gross Domestic Product (GDP)
- The actual growth rate of GDP
- The annual inflation rate
- Budget deficit

This index is calculated by assigning scores to different economic dimensions. The minimum score is zero, and the maximum score depends on the weight of each factor in the economy of the country. By combining these scores, a rank is assigned to countries regarding their economic risk. Higher values of this index suggest the greater economic risk of that country. As can be seen, in comparison to countries of the region, Iran has a high economic risk. Considering the direct and indirect effect of the banking system on the components of this index, improvement in the banking system situation of a country can be considered as one of the important factors for improving the economic risk index of that country.

Table 1
The economic risk index of the countries of the region (BMI Research, 2015)

Country	Index	Trend	Regional rank	Global rank
Israel	74.9	=	1	9
Saudi Arabia	70	=	2	26
United Arab Emirates	60.5	=	3	53
Qatar	59.4	=	4	56
Kuwait	59	+	5	59
Bahrain	58	=	6	64
Algeria	57.5	=	7	66
Oman	55.7	=	8	76
Egypt	55.2	=	9	80
Lebanon	55	-	10	81
Morocco	52.3	=	11	88
Iraq	49.3	=	12	94
Jordan	47.4	=	13	108
Tunisia	44.5	=	14	119
Iran	39.1	=	15	140
Libya	37.8	=	16	149
West Bank and Gaza	36.6	=	17	157
Syria	23.8	=	18	183
Yemen	21.4	-	19	186

Banking systems in different countries and especially in developing countries are always regarded as very important and a driving machine of the economy, whose consolidation leads to strengthened economies of countries against different types of shocks or impulses (Amini and Ashrafi, 2016). In recent years, the balance sheet of Iranian banks has experienced severe fluctuations in response to the instability of the macroeconomy in the real part as well as the frequent changes in the regulation rules in the monetary and banking section. One of the important changes that has occurred in the real part of the economy is the deep and sustained recession affecting production. In response to this profound recession, the assets of banks have experienced increased deferred claims and involvement of resources. Meanwhile, exchange rate shock, the incidence of sanctions in the foreign economy, and major embezzlements in several banks of the country have caused the financial statements of the banking system of the country to undergo severe income fluctuations (Heydari and Ahmadian, 2015).

Generally, it can be stated that there is no special method for preventing a financial crisis in an organization. The important point is that a set of significant elements that cause the incidence of financial crisis, in the long run, should be constantly monitored so that preventive measures could be taken appropriately (Koyuncugil and Ozgulbas, 2012).

The importance of the role of banks in the economy of the country, on the one hand, as well as other challenges ahead of the banks in the country on the other, have made precise monitoring of the performance of banks and identification of risks related to this industry indispensable. Taking care of the performance of banks is crucial for different groups, including investors, investees, policymakers, and shareholders. Each of these groups pays attention to the bank performance according to their special objectives and further evaluates it. Any measures taken for improving and enhancing the efficiency of the banking system will cause improvement in savings, investment, and allocation of resources, whereby potential, scattered, and hidden facilities in the country could be employed for development and public welfare. On the other hand, to prevent loss by the bank, decisions should be adopted for the management of banking risks; each of these decisions for risk management can have a positive or negative effect on the efficiency of a bank (Panahian and Abyak, 2013).

Today, in most advanced countries of the world, banks and financial institutes are evaluated based on various models and methods and ranked further. One of the indicators is the CAMELS index. This index is considered as a comprehensive assessment that captures all parameters and different aspects, including profitability, manageability, status among the competitors, etc., and as such is used in most advanced countries to rank financial institutes (Sudani, 2017). Nevertheless, the fact is that due to the changes in the banking paradigms of the world regarding various aspects such as financial, technological, and other dimensions, researchers believe that in a not very far future, assessment of banks through CAMELS method will not be efficient enough; instead, some methods should be developed to capture more items of financial statements and environmental factors (Babu and Visvanatan, 2018). Hence, novel methods should be assessed and financial risk of banks should be monitored.

This paper aims to present an EWS model based on data mining. This model has been presented to identify the risk profiles of the studied banks, to detect the most important influential ratios affecting the risk of banks, and to show the roadmap to improve the financial status of banks. The data mining method utilized in this paper is the chi-square automatic interaction detector (CHAID) decision tree algorithm. Accordingly, the results obtained from this research can first be utilized by financial researchers and specialists. Also, considering the easy understandability of the decision tree, it can even be used by non-specialist people in this area. Eventually, the obtained ranking has been compared with the results of the CAMELS model.

In comparison to other research conducted in the area of performance measurements of Iranian banks, this paper covers a broader spectrum of audience.

On the one hand, policymakers and regulatory institutions, by investigating the risk profile of banks, would be able to understand the general situation of the banking system and further identify the strong and weak points of the financial performance of banks. On the other hand, managers of banks would be able to achieve a proper understanding of the current situation of their bank as well as a roadmap for its improvement. Eventually, investors and stakeholders of these banks, by observing the financial performance profile of banks, can make better decisions.

In the second part of the paper, the theoretical background is presented. We deal with both domestic and foreign research conducted in the area of designing and implementing EWS for monitoring the financial performance of banks. The second part addresses the method of the research, whereby the model of interest is implemented. Next, data are analyzed, and the model is implemented. Finally, after presenting the experimental results of the research, the paper will be concluded.

2 Theoretical Background and Literature Review

2.1 Early Warning System (EWS)

EWS is a system used for predicting the level of success, possible abnormal events, and mitigation of risk of crisis in trades, systems, phenomena, companies, and people. Also, the current status and potential risk of the mentioned points can be calculated quantitatively (Koyuncugil and Ozgulbas, 2010).

A financial EWS is a regulatory and reporting system that warns possible financial problems and crises before their emergence in financial statements. These systems are employed for identifying financial efficiency, possible risks, and the potential bankruptcy of companies. Indeed, these systems, based on available probabilities, help the management avoid possible problems and crises.

The warning system was first propounded following the currency crisis of European countries in 1992 and 1993, the plight of Latin American countries in 1994 and 1995, and more seriously following the crisis of Eastern Asian countries in 1997 and 1998. In this regard, in addition to the International Monetary Fund, which is a pioneer of this method, some universities and central banks have also performed some research (Sayadnia et al., 2010).

Since the most important financial information of any organization is recorded in financial statements, hence the information of these statements is considered as the primary input for these warning systems. EWSs can be a subset of management information systems (MIS), decision support systems, and expert systems considering the method of design and the algorithms utilized.

Legislators and supervisors of the banking system can employ the EWS in at least the four following areas:

1) Investigating the operations and performance of the bank and discovering its violation of central bank laws
2) Studying the online banking system and assessing the electronic cash transfer
3) Finding the bank managers' and employees' violation of their legal duties and rights (fraud investigation methods)
4) Investigating the financial status of banks and confirming or denying the financial health of banks (Dikin, 2016).

In recent years, the attention of regulatory institutions, policymakers, and central banks of developed countries has increasingly been directed to EWS. For example, O'Brien and Wosser (2018), under the supervision of Ireland central bank, designed an EWS for financial crises of banks. In this research, the quarterly data 27 developed economies within 1980-2016 were used. This research generated a flexible EWS, which predicted the effects of the financial crisis 2008 eight seasons before its occurrence.

In particular, following the financial crisis in 2008, the European Union has emphasized the application of EWSs for preventing financial crises. In this regard, the EU has defined a project entitled "European early warning for systematic risk," so that policymakers and decision-makers could gain awareness about the incidence of crisis before its occurrence and take preventive measures based on signals they receive.

The present paper emphasizes "investigating the financial status of banks and confirming or denying the financial health of banks." Hence, using a data mining approach, an EWS is designed and used further for categorizing banks in terms of financial risk. In addition to allowing bank managers to better investigate the financial status and risk profile, this system can also be used by investors and shareholders of banks for financial risk assessment.

2.2 Data Mining and Its Application in Designing EWS

Data mining is an interdisciplinary subject specially covered by computer sciences. This field explains the computational process of discovering patterns in large data sets using artificial intelligence and machine learning methods as well as statistics and database systems. Indeed, data mining refers to implementing a set of methods with different orders such that some unavailable information, which is obtainable only through data analysis, could be extracted (Fesengheri et al., 2015).

Daffern data mining methods employ various subbranches of different sciences. These subbranches include a branch of science called exploratory data analysis as well as a branch of artificial intelligence science called knowledge discovery. Typically, the goal of data mining is to discover information that cannot be extracted from hidden and analytical dimensions of the sheer volume of data without complex statistical analyses as well as pattern extraction methods that are based on learning or without the extraction of analytical features. Indeed, as the volume of data grows and the relations between data become complicated, the discovery of analytical information across data becomes complex further, whereby the significance of data mining in extracting useful analyses becomes evident. Hand et al. also presented the following definition for data mining: "the science of extracting useful information from large amounts of data (2001)".

Indeed, data mining is a set of techniques for predictive extraction of hidden intimation among data. By predicting future trends and behaviors, this tool results in active and knowledge-based decision-making for an organization. Data mining is the science and technology of exploring through data for discovering unknown patterns in databases (Soon and Lee, 2007).

Technically, data mining refers to the semi-automatic process of knowledge extraction in the form of a pattern and from a set of general knowledge.

Data mining methods are categorized into two groups: supervised and unsupervised. The supervised methods seek a specific and predefined goal. In other words, they try to find a special pattern among the data available in the dataset. On the other hand, the reason for applying unsupervised methods is finding patterns or similarities among groups of data in which there is a predetermined set of groups or patterns. The important point regarding supervised data mining is the presence of a special target variable that should be categorized, predicted, or estimated. However, in unsupervised data mining, the task of data mining is to find general patterns among data, i.e., no particular pattern or target variable is intended.

The main steps of data mining can be described as follows:

- Retrieving data from a large dataset
- Choosing the subsets associated with the task of interest
- Decision-making regarding proper sapling system, clearance of data, and missed records.
- Implementing correct mapping on data
- Obtaining a model for the prepared data (Fayyad et al., 1996).

Considering the nature of EWS, there is a close relationship between the design of these systems and data mining techniques. It is because the main aim of EWS is first to purify and then to classify as well as analyze information for timely signal presentation. To this end, generally, unsupervised data mining techniques such as clustering are used for data preparation, while supervised methods such as decision trees are utilized for the final classification. Also, considering the EWS users, the application of techniques such as decision trees will help understand the analysis output more quickly. Usage of statistical methods and prediction techniques is another function of EWS, which will become possible through data mining.

The most common data mining techniques are as follows:

- Artificial neural network: Artificial neural network is an approach to data processing inspired by the biological nervous system and processes the information like the brain. The main point in this approach is the novel structure of the information processing system. This system benefits from a number of processing elements arranged in a specific structure, which practically provides various advantages for the user regarding information processing.

  Decision tree: Decision tree is one of the powerful and famous tools in data mining. Indeed, the decision tree is a data structure for dividing large recordsets into smaller sets. It occurs through a set of straightforward decision-making rules.
- Genetic algorithm: It is one of the metaheuristic algorithms. Genetic algorithm is one of the optimization techniques which employs simulation of the evolutionary process. The major components of this system include mutation, combination, and selection.
- The nearest neighbor method: It is a nonparametric method used for classification and regression problems, though it is used for classification problems most of the time.
- Association rules exploitation is one of the critical data mining methods. This method falls in the group of unsupervised learning methods and

intends to discover and extract local algorithms. Association rules represent the interdependence among data in databases.

In this paper, the decision tree has been used. A decision tree is a structure similar to a tree that reveals a set of decisions. This method presents a set of rules for data classification. Examples of decision tree algorithms include CART and CHAID. These algorithms create a set of practices employed for application on new unclassified datasets. The data preparation process is more straightforward in the CART algorithm than in CHAID (Lee and Siau, 2001).

The main objective of EWS discussed in this paper is to identify different risk levels as well as the factors affecting financial performance. CHAID algorithm deals with the classification of different groups according to the extent of connections based on chi-square criteria. Thus, the leaves of the CHAID decision tree do not have binary branches, and the number of branches will be at most equal to the number of variables. Modeling based on CHAID is an exploratory analytical method that is suitable for studying the relationship between the target variable and a large number of independent variables (Tirling, 2014). Thus, the use of this algorithm will be appropriate for this research. Note that usage of a decision tree for data classification reduces the prediction accuracy to some extent. Hence, based on the type of problem, a decision would be made about using or not using this technique.

2.3 CAMELS Model

CAMELS indices were first used in October 1987 by the National Credit Union Administration (NCUA) (Babar & Zeb, 2011). The Federal Reserve also evaluates its regulated banks using CAMELS indices, with each regulating one aspect of the bank financial health within a scale from 1 to 5. Rank 1 represents the best performance, while rank 5 shows the weakest performance. In this ranking, credit rate, profitability, and liquidity are among the most important criteria to determine the qualification and measure the activity of a bank. For this purpose, since 1988, the Basel Committee on Banking Supervision (BCBS) has considered the application of CAMELS indices as necessary for assessing financial institutes. The components of this index are as follows:

1) Capital adequacy: The bank would have to keep a reasonable level of capital. Basel committee has considered the minimum capital adequacy ratio for the banks of the Organization for Economic Cooperation and Development (OECD) member countries as 8%; however, for banks in Asian countries, due to weakness of credit monitoring system, the minimum ratio has been considered as 12%.

2) Quality of assets: Quality of assets in banks is directly associated with their financial performance. The value of loans is dependent on the value of liquidity of collaterals, while the value of investments is a function of the market value. A bank is expected to use stable assets in the portfolio and consider a scheduled plan for the reduction of the value of its assets as well as suitable reserves to compensate for this value.
3) Quality of management: management quality assessment represents the cost of every monetary unit given as a loan. Thus, its reduction leads to enhanced efficiency and profitability of financial institutions. The performance of the four other components of CAMELS is contingent upon the insights, abilities, awareness, processualism, perfectness, and qualification of managers of financial institutes. As the role of management is crucial in the success of any institution, generally, the quality management claims a lager weight compared to other indices of CAMELS elements in assessing financial institutes.
4) Revenues: the quality and trend of acquiring revenues in a financial Institute are closely related to the way the assets and liabilities are managed in that institute. Revenue acquisition in a financial institute should be profitable, such that it could support the growth of assets and increase the potential of saving in the organization to eventually enhance the shareholders' equity. An excellent revenue performance leads to increased reliability for investors, depositors, loaners, and the public sector towards the institute.
5) Liquidity: controlling liquidity is one of the important responsibilities of bank management. Usage of short-term sums in long-term investments causes a risk for the bank, whereby the owners of investment accounts may request to receive their sums. It compels the bank to sell its assets. A bank should have adequate liquidity to respond to the demand of depositors and loaners to attract general trust. Banks should have an effective asset and liability management to be able to minimize the mismatch between the usance in liabilities and assets and to optimize their return. Further, liquidity has an inverse relationship with profitability; therefore, financial institutes should establish a proper balance between liquidity and profitability (Abasgholipour, 2010).

2.4 Background of Foreign Literature

Attempts for classifying crisis-stricken companies were initiated by Beaver using the z-score method, who dealt with Altman analysis through financial ratios (Beaver, 1966). Over time, more research attention was paid to the supervision of the status of organizations for discovering dramatic changes and unexpected components of risk. Some of these studies are summarized as follows.

Tam (1991) using a back-propagation neural network, predicted the incidence of the financial crisis in the banking sector. Through probability investigation, he examined the bankruptcy of a set of Spanish banks in two forecast horizons: one year and two years. Also, Tam and Kiang (1992) investigated the predictive power of "linear auditing analysis," "logit," "nearest neighbor," "interactive binary separator," "feedforward neural network," and "backpropagation neural network" models in the area of banking crisis in terms of the bankruptcy of banks. Bell (1997) compared "backpropagation neural networks" with a logit model for predicting bankruptcy of the banking sector. He stated that in complex decision-making processes, the neural network model functions better. Olmeda and Fernandez (1997) examined the true predictive power for bankruptcy in the banking sector in the form of individual categorizing models and hybrid models (hybrid models consisting of several classifier models). Also, Alam et al. (2000) used fuzzy clustering and self-organizing map neural network models for classification to detect crises in the banking sector as a bankruptcy.

Kibritçioğlu (2004), in a paper entitled "analysis of the early warning signals of currency crises in Turkey from 1996 to 2004", tried to interpret the early warning system within a signal approach to identify the currency crisis, for which he used the exchange market index. By combining 46 economic variables, he could identify five critical periods in the economy of Turkey in 2000, among which 4 were related to the period before 2000, and one associated with 2001; his prediction was successful.

Brockett and Cooper (1990) employed a neural network method to design early warning systems. Initially, the model had been designed with 24 variables, but it eventually diminished to eight variables. These variables were shareholders' equity, capital ratio, return on assets, asset turnover, equity accounts receivable, loss, and liability changes to the current asset.

Lee & Urrutia (1996) compared logit as well as Hazard models, neural networks, and partitioned method to design early warning systems. For each model, they found different indices and concluded that the predictive power of all of the methods was the same.

Latinen & Chong (1999) employed EWS signals to design a system for predicting the crises of small companies. This research was a summary of the results of two different research in Finland and England. Both studies included seven major factors and around 40 sub-factors as the components of organizational failure. To express the significance of factors, the moving average method was used between the two studies. There were considerable similarities between them, where management inadequacy was found as the most important factor in both of them.

Yang et al. (2001) employed neural networks to design an EWS for discovering the financial risk of banks. They found that this method is a useful instrument for designing EWS.

Edison (2003) designed an operational EWS to discover financial crises. This system monitored numerous indices that could lead to a financial crisis; in case they exceeded a certain threshold, it issued a warning suggesting the possibility of incidence of financial crisis within the period of interest. The system was tested on the financial crisis of 1997 and 1998. In spite of some weak points, the results suggested that a quick warning system is a useful tool to discover risk and financial crises.

Jacobs and Kuper (2004) designed a warning system for six Asian countries. In this research, the financial crisis was defined as three factors: the crisis of liquidity, banking, and liability. The importance of these indices was tested in a multivariate logit model for six Asian countries within 1970-2001. The results indicated that the liquidity crisis was defined as a more serious problem compared to other crises.

Brockett et al. (2006) investigated the impact of statistical models and neural networks in discovering the financial problems of life insurers. They considered two neural network methods: backpropagation (BP) and learning vector quantization (LVQ). The results showed that these two methods outperform traditional statistical methods.

Abumostafa (2006) discovered warning signals for the exchange rate risk of Egypt, Jordan, and Turkey. The variables considered in this research were: real foreign exchange rate, exports, imports, gross domestic production, foreign assets, real domestic interest rate, global oil price, and the government's consumption. The results of that research indicated that all exchange rate crises were predictable, and this EWS proved to be efficacious.

Koyuncugil and Ozgulbas employed data mining to design a financial EWS for small and medium active companies in the Turkish market. They utilized the CHAID decision tree method to determine risk profiles and signs. This research covered 697 small and medium organizations within the timeframe of 2000-2005. In response to this research, 19 risk profiles were created, according to which 430 companies had a poor financial performance.

Tanaka et al. designed an EWS to predict the banking crisis based on the random forest method. To this end, they employed 48 indices, which could be categorized into four groups: profitability ratios, investment, quality of the loan, and financial resources. Investigation of data was performed within 1986-2014. They found that this system has a better performance in predicting the financial crisis of banks compared to other warning systems.

Davood et al. (2017) dealt with designing an EWS for predicting the government's liability crisis. The methods utilized in this paper included the Dynamic Signal Extraction and logit method. They found that this method outperformed other methods used for predicting the government's liability crisis.

2.5 Background of Domestic Literature

This section deals with investigating studies related to the subject in Iran. In this regard, Naderi (2003) addressed an EWS for predicting the incidence of the financial crisis (currency crisis) in Iran's economy from 1959 to 1999 using the sign extraction method and logit model. In this research, first using the signal method, the important indices affecting financial crises, and then by presenting a probability function for the financial crisis through the logit model, the critical years in Iran's economy were identified. Kanani (2005) addressed the prediction of the incidence of a currency crisis in oil-dependent economies including Iran. In this paper, the role of variables such as changes in the real exchange rate, changes in the volume of exchange reserves, variations in monetary variables in relation to the exchange reserves was investigated with regards to the prediction of currency crises alongside oil shocks in prominent countries in terms of oil resources. They concluded that if oil price changes, exchange reserve changes, and variations of the cash volume to the currency reserve volume ratio issued a warning concurrently, the probability of incidence of crisis would reach 100%. Erfani (2006) developed an EWS for the monetary crisis in Iran’s economy using a Markov Switching Model for the time period 1988-2004. Sayadnia et al. (2010) employed a signaling method, logit model, and neural network to present a warning system to identify financial crises in Iran's economy. Shajari and Mohebikhah (2010) utilized a signaling method and developed a probability model for predicting the incidence of banking crisis and crisis in the balance of payments (twin crisis). Based on the results obtained from this research, two variables of stock prices and real interest rates were the most credible indices for predicting the monetary crisis. Also, the stock prices alongside the real exchange rate were found as the most suitable indices for predicting monetary crisis as well as twin crises. Abounouri and Erfani (2008) employed a Markov switching model and investigated the annual data of OPEC member countries including Iran from 1989 to 2003. They also designed a financial EWS for them. Moshiri and Nadali (2010) identified banking crisis years in Iran's economy within 1961-2008 using Markov's simple switching model and Markov switching GARCH model and based on the exchange market pressure index. Ghavam et al. (2015) determined past years of Iran's economy during 1978-2012 based on investigating the behavior of rate of changes in the macroeconomy variables in Iran within this period in the form of four indices including standard deviation, moving standard deviation, semi-deviation, and moving semi-deviation followed by "self-organizing map" neural network model. In the next stage, using the feedforward neural network model, they predicted the future values of the rate of changes in the mentioned variables between 2013 and 2016. Eventually, based on the obtained results and using the "pattern network," neural network, criticality or non-criticality of the mentioned years was predicted. The final results of this research indicated that the financial crisis in Iran in 2012 had its roots in 2011 and 2010. In spite of its persistence in 2013, this crisis gradually disappeared during this year, with 2014 and 2016 being non-critical years.

Nasrollahi et al. (2017) employed seasonal data of Iran's economy within 1988-2014 and used a model with a discrete dependent variable. In addition to investigating the factors affecting the incidence of a currency crisis in Iran, they designed and interpreted an EWS for currency crises with all required elements regarding Iran's economy. They concluded that the designed system can mostly explain the determining factors of a currency crisis in Iran, and has had a great ability in predicting these crises within the tested periods. Based on the obtained results, currency crises in Iran are a result of different imbalances in real and public sectors, foreign balance, and financial sector of the country continuously. Based on the results of this research, the variables of loan to deposit ratio, the banks' debt to the central bank to the monetary base ratio, inflation rate, and developed industrial production have had major roles in increasing the probability of development of currency crises in Iran.

3 Statistical Population and Sample Selection

The population of interest consisted of all institutional and credit institutes as well as banks regulated by the Central Bank of the Islamic Republic of Iran (34 banks and credit institutes). Since national banks are not a member of the stock market, because of the unreliability of their financial reports, they were excluded from the research. Further, since in investigating the annual financial statements of some banks and financial institutes due to improper disclosure of information, it was not possible to extract and calculate research variables, hence they were also excluded from the research to keep homogeneity. Eventually, based on the above points, 13 banks and financial institutes were chosen for the study.

4 The Subject, Spatial, and Temporal Scope of the Research

The subject scope of this research is assessing the application of a financial EWS in the Iranian banking system with a data mining approach. In this research, the data of banks, as well as financial and credit institutes regulated by the central bank of the Islamic Republic of Iran, were used, and hence it does not include the banks and institutes of other countries.

The initial data were collected in January 2019, and preliminary analysis was performed. In July 2019, as the time range studied increased, new data were added to the research and analyzed further.

5 The Methodology of Implementation of the Model

The EWS model designed using data mining can be seen in Fig. 1. EWS stages are as follows (Koyuncugil and Ozgulbas, 2012):

1) Data preparation
2) Implementation of data mining technique
3) Determination of the risk profiles
4) Specification of the current status of banks based on risk profiles and determined risk indices
5) Definition of a roadmap for improving the status of banks

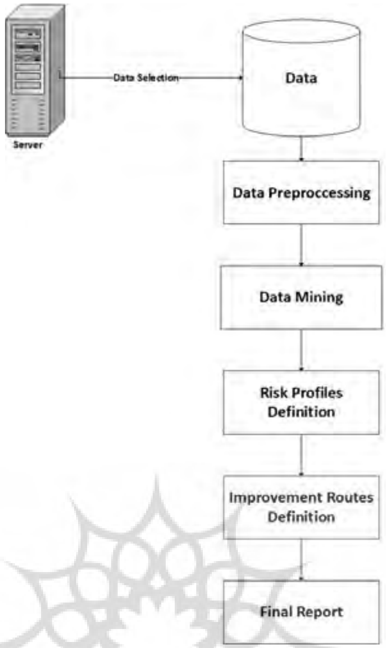

Figure 1. The EWS Data Flow Model

Step 1: Data preparation

Financial data were extracted from financial statements of the studied banks and utilized for calculating the financial performance indices (according to Table 2). Note that these ratios have been adopted from Koyuncugil and Ozgulbas's (2012) research. The data have been collected from the financial statements of the studied banks from 2003 to 2017, including 122 financial statements. The data from 2003-2014, including 83 financial statements have been used as the data of the training stage for the decision tree, while the data from 2015 to 2017, including 39 financial statements have been employed as test data. At this step, the following stages have been taken:

1) Calculating the financial performance indices according to Table 2
2) Reducing the indices to resolve the correlation/collinearity problem: for this purpose, the pairwise correlation coefficient of variables was calculated through the Pearson correlation coefficient method. The variables with a correlation coefficient above 0.8 were chosen for reduction.
3) Substituting missing data

4) Removing outliers

Table 2
The financial indices of the research

Definition	Financial variables	Symbol
Net profit to sales	Net Profit Margin	VAR1
Operational profit to sales	perational Profit Margin	VAR2
Net profit to the sum of assets	Return on assets	VAR3
Net profit to capital	Return on investment	VAR4
Net profit to working capital	Return on working capital	VAR5
Equity to return on assets	Loan profitability measurement	VAR6
Current assets to current debts	Current ratio	VAR7
Assets apart from the stock of materials and goods to the current debt	Quick ratio	VAR8
The cash flow of operations to the dividend paid, purchase of a fixed asset and long-term debt reimbursements	Liquidity Coverage Ratio	VAR9
Cash flow resulting from operations to the current debts	Cash cycle ratio	VAR10
Current assets to total assets	Current asset ratio	VAR11
The cost of production to the book value of fixed assets	Total asset turnover	VAR12
Sum of debts to the sum of assets	Debt ratio	VAR13
Sum of debts to the shareholders' equity	Debt to equity ratio	VAR14
Sum of long-term debts to the shareholders' equity	Long-term debt to equity ratio	VAR15
Sum of current debts to the shareholders' equity	Current debt to equity ratio	VAR16
shareholders' equity to the sum of debts	shareholders' equity to sum of debts	VAR17
shareholders' equity to the sum of granted facilities	shareholders' equity to facilities	VAR18
Sum of assets to the shareholders' equity	assets to the shareholders' equity	VAR19
Facilities paid to assets	Facilities paid to assets	VAR20

Table 3
The profiles obtained from CHAID algorithm

Var18	Var15	Var9			
1	**0, 3**	<= 11.56			
2	0, 2, 7	11.56 < 35.894=>	> 0.168		
3	0, 1, 4	>35.894		> 0.03	
4	0, 1, 5	>35.894		<=0.03	
5	0, 2, 6, 10	11.056< 35.894 =>	<=0.168	> 0.02	
6	0, 2, 6, 9	11.056 < 35.894 =>	<=0.168	0.01 < 0.02 =>	
7	0, 2, 6, 8, 11	11.056 < 35.894 =>	<=0.168	=< 0.01	=< -9.34
8	0, 2, 6, 8, 12	11.056 < 35.894 =>	<=0.168	=< 0.01	> -9.34
		Var19	Var18	Var15	Var9

Step 2: Implementing data mining method (CHAID)

One of the objectives of designing EWS is to create a system whose application is simple for those with no special financial or data mining knowledge, and they can quickly investigate the current status of the company and make the decision accordingly.

Thus, from among the mentioned methods, the decision tree method is suitable because of its simple understanding and graphical representation of the results. From among different algorithms presented so far for the decision tree, CHAID is superior to other algorithms of decision tree because of the number of branches it can create. This algorithm is developed based on two groups of variables called target variable and predictor variables. In this paper, the calculated financial ratios represent the financial performance of banks. Hence, the variable "bank financial performance" with three values of "good," "medium," and "poor" have been considered as the target variable, while financial ratios (reported in Table 2) have been regarded as the predictor variable. Fig. 2 demonstrates the CHAID decision tree and its resulting profiles. Table 3 outlines the rules created by the decision tree and the method of allocation of every financial statement to the financial performance profiles. Also, the figures related to the number and percentage of purity of each node can be seen based on the test data. Note that in the design decision tree, the bank financial performance profile refers to a node with no sub-branch. As observed in Fig. 2, the obtained decision tree has eventually ended in eight profiles.

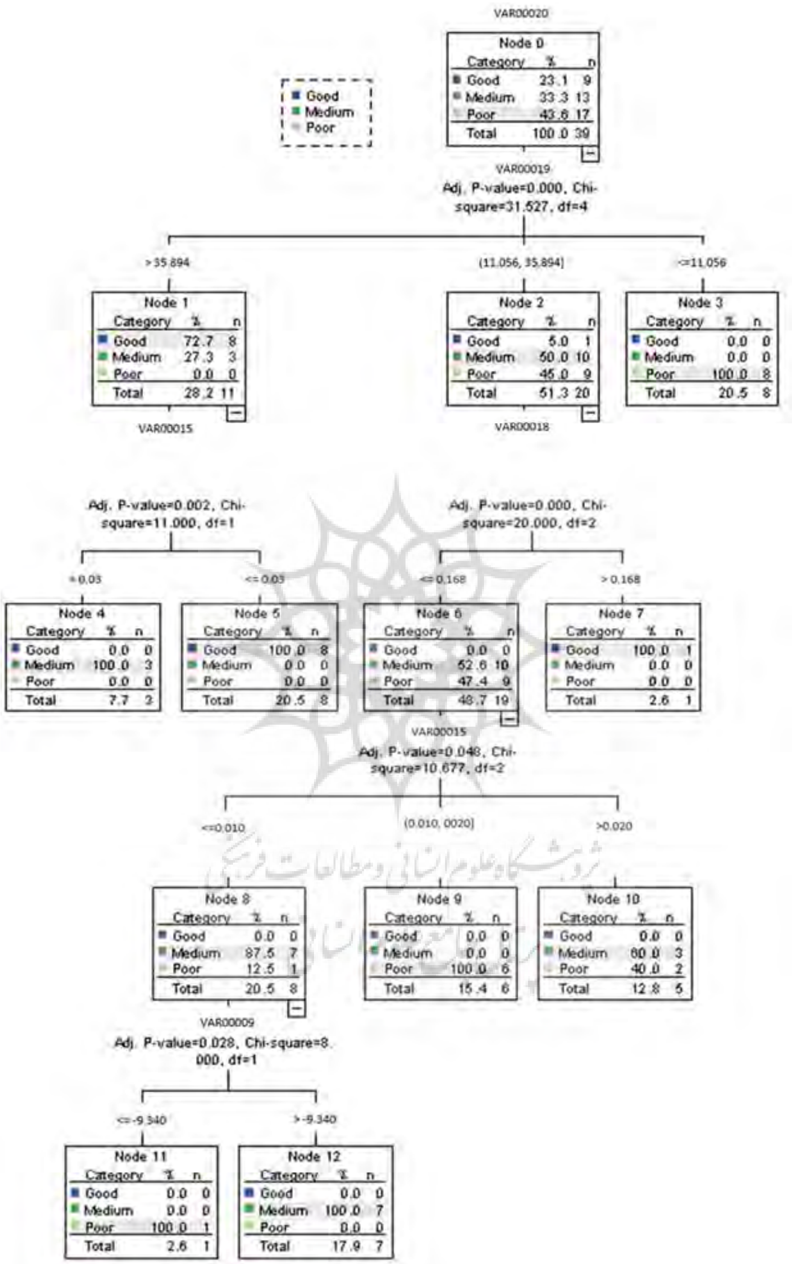

Figure 2. CHAID Decision Tree and the Profile of Financial Status of Banks.

Step 3: Determining the financial performance profiles

Since the CHAID algorithm creates a multi-branch decision tree, significant relationships between predictor variables and the target variables can be well-identified. As seen in Fig. 2, although the financial performance of banks can be easily categorized only into three groups of good, medium, and poor, based on the created tree, in addition to the mentioned classification, another more detailed and precise categorization is also possible based on the financial performance profiles and comparison of the performance of financial statements of each profile. The resulting tree indicates to which profile every bank belongs to based on the current status, and on what variables they should focus on improving their financial status. It is also possible for bank managers to choose a pattern for the desired state of the bank based on the best profile available. CHAID algorithm conducts a chi-square independence test between the target variable and predictor variables. The creation of branches begins from the variable showing the greatest correlation with the target variable, which continues accordingly.

Table 3 reports the status of financial statements for the last three years (test data) in the obtained profiles. Overall, 9, 13, and 17 statements are good, medium, and poor, respectively, claiming 23.07, 33.33, and 43.6% of the data.

Table 4

The Financial Status of the Studied Banks

Profile	**Financial performance**					
	Good		Average		Poor	
	No.	%	No.	%	No.	%
1					8	100
2	1	100				
3			3	100		
4	8	100				
5			3	60	2	40
6					6	100
7					1	100
8			7	100		
Total	9	23.07	13	33.33	17	43.6

As can be seen in Fig. 2, the financial performance of the organization (VAR20) has the highest correlation with the asset to shareholders' equity ratio (VAR19) (p-value=0). The banks with assets to shareholders' equity ratio less than or equal to 11.056 falls in profile 1. This profile, with a purity of 100%, suggests the poor financial status of the bank. In case this ratio is between 11.056 and 35.894, the bank lies in node 2; the banks in this node are categorized into two groups based on the shareholders' equity to loans ratio (VAR18): Group 1 covers the banks in which this ratio is larger than 0.168, which are assigned to profile 2 and have a good status; on the other hand group two covers the banks in which this ratio is less than or equal to 0.168, and are assigned to node 6. The banks of node 6, based on the long-term debt to equity ratio, lie in either profile 5, or 6, or 8. The banks in profile 5 have a medium or poor status. This profile is the only profile obtained from the decision tree whose purity is not 100%. Therefore, the error of the decision tree can be calculated. In total, out of the 39 data used for training the decision tree, there are two errors, meaning that the efficiency of this decision tree is 94.8%.

Next, in case the long-term debt to equity ratio is larger than 0.01 and smaller than or equal to 0.02, the bank will be in profile 6. The banks in this profile have a poor financial status. Further, if this ratio is smaller than or equal to 0.01, the judgment is made based on the liquidity coverage ratio. In case this ratio is greater than 9.340, the bank will be in profile 8 and has a medium status. On the other hand, if this ratio is smaller than or equal to 9.340, it belongs to profile 7 and will have a poor financial status.

Step 4: Determining the current status of the bank

So far, the risk profiles were identified based on the available data. For this purpose, the historical data of banks were used. At this stage, every bank achieves the current status of the organization by comparing its key variables against the decision tree. Table 5 reports the variables utilized in the decision tree. Based on the results of the research, these variables affect the target variable, i.e., bank financial performance. Considering the performance profiles and variables, two ratios of an asset to shareholders' equity and shareholders' equity to loans can be regarded as the signaling variables for the financial performance of banks. The general rules that banks can consider as a signal are as follows:

- If the asset to shareholders' equity ratio is less than 35.894, this means that the bank's financial status is poor.
- If the shareholders' equity to loan ratio is less than or equal to 0.168, this demonstrates that the financial status of the bank is poor.

Table 5
The variables affecting the financial status of banks based on the decision tree

P-Value	Description	Variable
0.0005	Asset to equity ratio	Var19
0.0006	Equity to facilities	Var18
0.048	Long-term debt to equity ratio	Var15
0.028	Liquidity coverage ratio	Var9

Based on the latest financial statements published by the 13 studied banks in this paper, the current status of these banks has been determined with its results reported in Table 6. The results suggest that four banks have poor while nine banks have medium financial status. Further, the largest number of banks have belonged to profile 5 claiming 46% of the share.

Step 5: Defining a roadmap for the banks

By identifying the best profile, it can be used as a model for achieving the roadmap. Banks can take steps for improving conditions based on their current status and the designed roadmap and according to the variables affecting their financial performance. Based on the decision tree in Fig. 3, the best risk profiles are profile 2 and profile 4. Thus, every bank should strive to obtain these profiles. The best roadmap for achieving these two profiles is presented in Table 7.

Table 6
The financial performance profile of banks based on the CHAID decision tree

No. of bank	Asset to equity	Equity to facilities	Long-term debt to equity	Liquidity coverage ratio	Profile
1	11.91	0.12	0.04	-12.38	5
2	22.97	0.06	0.09	-6.36	5
3	25.75	0.06	0.07	-12.45	5
4	14.18	0.09	0.08	-15.16	5
5	27.38	0.08	0.06	-10.35	5
6	13.78	0.1	0.01	-51.52	7
7	59.41	0.02	0.06	2.5	3
8	21.47	0.07	0.13	-3.57	5
9	6.56	0.26	0.01	0.48	1
10	-6.44	-0.27	0	-4.01	1
11	74.16	0.02	0.04	-25.56	3
12	10.27	0.13	0.01	-1.1	1
13	76.89	0.02	0.04	0	3

Table 7
The roadmap for achieving the desired profile

Influential variables			Profiles	Roadmap
Long-term debt to equity	shareholders' equity to facilities	Asset to shareholders' equity		
	0.168 <	(11.056, 35.894]	2	1
0.03 >=		35.894 <	4	2

Hence, the proposed method suggests that every bank, to enhance its asset to shareholders' equity ratio, should target at least 11.056, increase the shareholders' equity to loans to more than 0.168, and reduce the long-term debt to the equity ratio to less than 0.03.

6 Comparison of the Results with the Ranking of Banks Using the CAMELS Method

This section compares the results of the research with the results of the classification performed by the CAMELS method.

Erzae and Ghasempour (2017), in their paper entitled "ranking Iranian private banks based on CAMELS model using a hybrid approach of analytical hierarchy process and ARAS," evaluated Iranian private banks. Table 8 compares the output of the decision tree and CAMELS model.

The results suggest that the financial performance profile developed by the decision tree is, to a large extent congruent with the rank obtained from the CAMELS model. The only difference might be observed for bank No. 6; with the ranking of 5, it is expected that it would be categorized as low or medium risk bank. The superiority of the decision tree method over the CAMELS method is that to design the decision tree of interest using the CHAID algorithm, the number of utilized variables can be increased so that the most relevant variables would be chosen. Further, after the creation of the decision tree, decision making regarding the financial risk status of the bank becomes limited to investigating five variables, which is far simpler and faster as well as more efficient in comparison to employing the CAMELS model for decision-making. Finally, the decision tree helps in determining the roadmap for improving the bank status, while the CAMELS model only describes the current state of the bank.

Table 8
Comparison of the financial status of banks in the decision tree and ranks based on CAMELS model

No. of bank	Results of CHAID decision tree		
	Decision tree profile	Financial performance	Rank based on CAMELS model
1	5	Medium	6
2	5	Medium	9
3	5	Medium	8
4	5	Medium	7
5	5	Medium	4
6	7	Poor	5
7	3	Poor	13
8	5	Medium	2
9	1	Poor	12
10	1	Poor	10
11	3	Medium	3
12	1	Poor	11
13	3	Medium	1

7 Conclusion

A financial EWS is an analytical method for the current status of the bank in terms of financial performance and helps us in predicting the future status of the databank and reducing the risk resulting from unfavorable financial performance. By applying this analytical technique, the potential risks arising from improper financial ratios are identified, and preventive measures can be taken before the incidence of the financial crisis. The advantages of employing the EWS systems for banks can be summarized as follows:

- Warning the bank before the incidence of crisis to take preventive measures
- interpreting the roadmap to achieve the desired status
- facilitating the decision-making process
- helping in monitoring the extent of objectives' accomplishment

The development and application of operational solutions will support both banks and the economy of the country further. Possessing information about financial performance, monitoring financial performance, and achieving the roadmap is essential to improve the current status for every bank. In this regard, data mining as a reflection of information technology in the area of strategic decision-making presents a system for promoting the decision-making power of bank managers in the field of financial performance in the form of a warning system. Thus, the application of these systems by banks first helps the managers better understand the current status. It also leads to faster and facilitate decision-making, and can even change into a competitive advantage for the bank. Besides, by applying this warning system, regulatory organizations can also obtain a proper yet simple classification of banks.

In this paper, a financial EWS was presented based on data mining for assessing the financial performance of several private banks in the Iranian banking system. According to the results, 83 financial statements were used to design the CHAID decision tree; based on the designed tree, 39 financial statements were categorized. The results suggested that out of the 39 classified financial statements, 23.1, 33.3, and 43.6% had good, medium, and poor performance, respectively. The banks lying in profiles 1, 6, and 7 had poor performance, those in profiles 3, 5, and 8 had media performance, and banks in profiles 2 and 4 showed good performance. Based on these profiles, according to the latest reports published by the studied banks, eight banks had medium and five banks had a poor financial performance.

Based on these profiles, four variables of an asset to shareholders' equity, shareholders' equity to loans, long-term debt to equity ratio, and liquidity coverage ratio were identified as the most relevant variables associated with the financial performance of banks.

In addition, a roadmap was presented to help banks achieve the desired status. The two roadmaps presented emphasized increasing the asset to shareholders' equity ratio to at least 11.056, increasing the shareholders' equity to loans ratio to 0.168, and reducing the long-term debt-to-equity ratio to less than 0.03, depending on the profile each bank currently belongs to.

Eventually, the results obtained from the CHAID decision tree were compared with CAMELS model findings. Classification of the financial performance of banks based on the decision tree matched the CAMELS model results. Nevertheless, in the decision tree method, decision-making is faster and easier by investigating fewer variables.

References

Abasgholipour, M. (2010). Factors Affecting the Improvement of the Performance of Banks. *Banking and Economy Quarterly*, 106, 24-35.

Abounouri, A., Erfani, A. (2008). Markov Switching Algorithm and Predicting the Probability of Incidence of Liquidity Ratio in OPEC Member Countries. *Bulletin of Economics*, 8, 153-174.

Alam, P., Booth, D., Lee, K., & Thordarson, T. (2000). The Use of Fuzzy Clustering Algorithm and Self-Organizing Neural Networks for Identifying Potentially

Failing Banks: An Experimental Study. *Expert Systems with Applications*, vol. 18, 185-99.

Amini, A., Ashrafi, A. (2016). Investigating the status of the banking system in resistive economy emphasizing Islamic banking. Proceedings of the first congress on the resistive economy.

Babar, H. Z., & Zeb, G. (2011). CAMEL Rating System for the Banking Industry in Pakistan. *Master Thesis.Umea School of business*.

Beaver, W. (1966). Financial Ratio as Predictors of Failure. *Journal of Accounting Research*.

Bell, T. B. (1997). Neural Nets or the Logit Model? A Comparison of Each Model's Ability to Predict Commercial Bank Failures. *International Journal of Intelligent Systems in Accounting, Finance and Management*, vol.6, 249-64.

Brockett, P. L, & Cooper, W. W. (1990). Report to the State Auditor and the State Board of Insurance on Early Warning Systems to Monitor the Performance of Insurance Companies in Texas.

Brockett, P.L., L.L Golden, J. Jang, & C. Yang. (2006). A comparison of neural network, Statistical methods and variable. *Journal of Risk and Insurance*, 397-419.

Edison, H. J. (2003). Do Indicators of Financial Crises Work? An Evaluation of an Early Warning System. *International Journal of Government Financial Management*, 11, 53.

Erfani, A. (2006). The Currency Crisis and Iran's Economy: An Early Warning System, *Doctoral thesis*, Faculty of Humanities, Mazandaran University.

Erzae, A., & Ghasempour, S. (2017). Ranking of Private Iranian Banks Based on CAMELS Model Through the Hybrid Approach of Analytical Hierarchy Process And ARAS. *Financial Management Strategy*, 5th year, 18.

Fayyad, G., Piatetsky-Shairo, P., & Symth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine, 17(3), 37-50*.

Fesengheri, M., Bahrpeyma, F., Cheraghchi, H., & Abdollahi, A. (2015). *Data Mining Methods in the Stock Market*. First publication. Bourse Publications, Tehran.

Ghavam, M. H., Ebadi, J., & Mohammadi, S. (2015). Designing a Hybrid Model of Financial Crisis EWS for Iran's Economy. *Quarterly Research of Applied Economic Studies in Iran*, fourth year, 13.

Hand, D., Mannila, H., & Smyth, P. (2001). *Principle of Data Mining*. Cambridge: MIT Press.

Heydari, H., Ahmadian, A. (2015). Investigating the Impacts of the First and Second Round of Macroeconomy Transformations on the Financial Statement of Banks. *Monetary-Banking Research Quarterly*, eighth year, 23.

Jacobs, L. J., & Kuper, G. H. (2004). Indicators Of Financial Crises Do Work! An Early Warning System for Six Asian Countries. *CCSO Working paper* 13. Department of Economics, University of Groningen, The Netherlands.

Kanani, A. (2005). Predicting Currency Crises in Oil-Dependent Economies Using KLR Model, *MA thesis*, the Institute for education and research on management and planning, management and planning organization.

Kibritçioğlu, A. (2004). *An Analysis of Early Warning Signals of Currency Crises in Turkey*. 1986-2004. from www.kibritcioglu.com.

Koyuncugil, A. S., & Ozgulbas, N. (2010). Social Aid Fraud Detection System and Poverty Map Model Suggestion Based on Data Mining for Social Risk Mitigation. Surveillance Technologies and Early warning systems: Data Mining Application for Risk Detection. 173-193. New York: Idea Group, Inc.

Koyuncugil, A. S., & Ozgulbas, N. (2012). Financial Early Warning System Model and Data Mining Application for Risk Detection. Expert Systems with Applications.

Latinen, K., & Chong, H. G. (1999). Early warning system for the crisis in SMEs. Preliminary evidence from Finland and the UK. Small Business and Enterprise Development.

Lee, S. H., & Urrutia, J. L. (1996). Analysis of Insolvency Prediction in the Property-Liability Insurance Industry: A Comparison of Logit and Hazard Models. *Risk and Insurance, 63(1), 131-130*.

Lee, S. J., & Siau, K. (2001). A Review of Data Mining Techniques. Industrial Management & Data Systems, 101(1), 41-66

Moshiri, S., & Nadali, M. (2010). Identifying Banking Crises in Iran's Economy. *Economic Policies*, 78, 59-78.

Naderi, M. (2003). Presenting an EWS for the Financial Crisis for Iran's Economy. *Economic Research Quarterly*, 17, 147-154.

Nasrollahi, M., Yavari, K., Najarzadeh, R., & Mehregan, N., (2017). Designing an EWS for Currency Crises in Iran: Logistic Regression Approach. *Economic Research*. 52, 1.

O'Brien, M., & Wosser, M. (2018). An Early Warning System for Systemic Banking Crises: A Robust Model Specification. Central bank if Ireland.

Olmeda, I., & Fernandez, E. (1997). Hybrid Classifiers for Financial Multicriteria Decision Making: The Case of Bankruptcy Prediction. *Computational Economics*, vol.10, 317-35.

Panahian, H., & Abyak, H. (2013). Interpreting the Effects of Risk on the Efficiency of Banks through Calculating Efficiency via DEA Method. *Accounting and Auditing Research, 2013, 17, 23-45*.

Sayadnia Tayebi, A., Arshadi, A., Samadi, S., & Shajari, H. (2010). Interpreting a Warning System for Identifying Financial Crises in Iran. *Money and Economy Quarterly*, 6.

Shajari, P., & Mohebikhah, B. (2010). Predicting Banking Crises and Balance Of Payments through KLR Signaling (case study Iran). *Money and Economy Quarterly*, 4, 115-152.

Shankar Babu, M., & Viswanthan, E. (2018). Can the CAMELS Rating System Survive the Future? *International Journal of Scientific and Research Publications*, Volume 8, Issue 8, 355-365.

Soon, L. K., Lee, S. H. (2007). Explorative Data Mining on Stock Data Experimental Results and Findings. Printer ADMA, 562-569.

Sotudehnia, S., & Abedi, F. (2013). *The Impact of Monetary and Financial Policies on Financial Stability in Iran*, first-year, 3.

Sudani, A. (2017). Ranking Banks and Financial Institutes Based on CAMELS International Indices. *Monetary-Banking Research Quarterly*, tenth year, 31.

Tam, K. Y. (1991). Neural Network Models and the Prediction of Bank Bankruptcy. *Omega: International Journal of Management Science*, Vol.10, 169-86.

Tam, K. Y., & Kiang, M. (1992). Predicting Bank Failures: A Natural Network Approach. *Decision Science*, vol.23, 926-47.

Tanaka, K., Kinkyo, T., Hamori, S. (2016). Random Forest-Based Early Warning System for Bank Failures. *Economic Letters*, 118-121.

Yang, B., Ling, X. L., Jing, X. (2001). An Early Warning System for Loan Risk Assessment Using Artificial Neural Networks. *Knowledge-Based Systems*, 303-306.