academicJournals

Vol. 10(30), pp. 2902-2910, 23 July, 2015
DOI: 10.5897/AJAR2015.9883
Article Number: DDF447F54364
ISSN 1991-637X

http://www.academicjournals.org/AJAR

African Journal of Agricultural Research

Full Length Research Paper

Evaluation of insolvency in mutual credit unions by application of the data mining using decision trees approach

Isabel Cristina Gozer[1]*, Antônio Roberto Pereira Leite de Albuquerque[2], Sadao Isotani[3], Régio Márcio Toesca Gimenes[1], Willian Heitor Moreira[4], Odair Alberton[1,5] and Emilio Araújo Menezes[6]

[1]Paranaense University – UNIPAR, Umuarama, Paraná, Brazil.
[2]Departamento de informática em saúde, Universidade Federal de São Paulo - Escola Paulista de Medicina – UNIFESP-EPM, São Paulo, Brazil.
[3]Instituto de Física da Universidade Estadual de São Paulo – IFUSP, São Paulo, Brazil.
[4]MBA em Controladoria Gestão Empresarial e Financeira – Unipar, Umuarama, Paraná, Brazil.
[5]Postgraduate Program in Biotechnology Applied to Agriculture; Paranaense University – UNIPAR, Umuarama, Paraná, Brazil.
[6]Universidade Federal de Santa Catarina, Santa Catarina, Brazil.

Received 6 May, 2015; Accepted 13 July, 2015

This study aimed to present an evaluation of the insolvency of mutual credit unions in the Paraná State (Brazil) by application of the data mining using decision trees approach. The information required to build the models were obtained from indicators applied to a sample of 62 mutual credit unions from which 31 are solvent and 31 are insolvent. The selection of indicators was made based on the PEARLS system, whose efficacy refers to the World Council of Credit Unions (WOCCU). The decision trees were built by training the J48, ADTree and LADTree algorithms. After the analysis of results, the best performance was observed for the ADTree algorithm. According to the Kappa statistics, its acceptance level was excellent. In addition to the evaluation of performance of the decision trees, the paths with the highest confidence levels for assessing insolvency was identified by the A3 indicator (Net Institutional and Transitory Capital + Non-Interest-bearing Liabilities/ Non-earning Assets) (> 0.052), this value indicate that the cooperative is solvent. The confidence level was set at 1.953 and the path is represented on the second node of the tree.

Key words: Insolvency, credit unions, data mining, decision trees.

INTRODUCTION

The concern of economic agents in measuring the soundness of financial institutions, both individually and jointly, has increased with the current economic scenario influenced by the crisis of the real estate sector that

*Corresponding author. E-mail: isa@unipar.br, Tel: +55 (44) 36212828. Fax: +55 (44) 36212830.

began in 2007 in the USA. Within this scenario, the services of credit unions have grown significantly with lower interest rates and service costs.

Currently, the credit cooperatives are one of the major significant tools of social development since they favor credit democratization and income decentralization. Considering the current pattern of the National Financial System (Brazil), credit unions are an alternative to make credit accessible and include more people in the financial market by benefiting particularly those small and rural entrepreneurs with greater funds at lower interest rates than the average charged on the market (Oliveira, 2004; Braga et al., 2006; Gozer et al., 2014).

According to Chaves (2009), because of the constant uncertainties present in the Brazilian economy, commercial banks prefer to perform their profits on interbank transactions rather than on loans, thus creating a gap that should be filled by the State. However, the restructuring of the financial system at the beginning of the Real Plan eliminated several public banks and private financial institutions of small size that had a strong regional presence, which damaged the credit availability and aggravated regional disparities. Because this situation, discussions on the credit unions operations emerged to resolve the promotion of small entrepreneurs, thus filling the gap left by the market.

According to Chaves (2009), the credit cooperatives are strengthening the least developed regions of the country by reducing some disparities of the Brazilian economy. They play an important role in settling the difficulty of raising funds of small-sized companies and business. Quantitative models for predicting insolvency are built with basis on indexes commonly extracted from the financial statements of a sample of solvent and insolvent companies, from which characteristics are distinguishable. The final result of these models is a mathematic formula able to predict the future situation of a company, within a given margin of error (Gozer et al., 2014).

According to Bressan et al. (2011), one of the major challenges of the credit cooperatives is the creation of management models that meet all the peculiarities of the sector as well as the requirements constituting the doctrinal principles of the regulatory body, which is in the case of credit unions, the BACEN – Central Bank of Brazil.

Decision trees are an efficient way to develop classifiers from data mining, it is widely used due to its efficiency in what concerns the processing time and for putting up an intuitive way to analyze the results. They also present a form of simple symbolic representation and normally very comprehensible, thus facilitating the analysis of the problem (Garcia, 2003). From this context, this study aimed to assess the state of insolvency of mutual credit unions in the Paraná State (Brazil) by application of the data mining using decision trees approach.

THEORETICAL BACKGROUNDS

Trying to study efficiency in the cooperative credit sector has led to adopt new technology and managerial knowhow. Among the tools that facilitate efficiency, data mining has stood out in recent years as a sophisticated methodology to search for knowledge that is “hidden” in organizations' databases (Sousa and Figueiredo, 2014).

The decision tree is a technique classified in the context of data mining, used for sorting and generating data patterns. The knowledge is generated in a decision tree format that can be subsequently translated into rules. It is visually represented as a tree – the leaves are the rules with the classification of the analyzed data, which makes the data interpretation much easier for users. Each rule starts at the root of the tree and goes toward the leaves (Lima, 2007).

According to Han and Kamber (2006), data mining is the process of building knowledge from a great volume of information stored in a database and it can be considered the most important step in the process of acquiring knowledge. The data mining processes search by patterns, associations, changes, anomalies and significant structures among data, therefore, they can raise valuable information in large database (Islam and Habib, 2015).

According to Lemos (2003), Ross Quinlan, professor at the University of Sydney (Australia), was the creator of the technology that allowed for decision trees. Quinlan, as “the father of the decision trees”, developed the new ID3 algorithm by 1983. Both the ID3 and its evolutions (ID4, ID6, C4.5, See 5) are very well adapted for decision trees as they produce rules ordered by importance, which are used to build a decision tree model of the events affecting the output items. Decision trees are part of the classification methods and are always used with the technology of induction of rules. However, they are unique in the sense of presenting the results in a form of prioritization. Thus, the most important attribute of a decision tree appear in the first node, and the attributes of decreasing relevance are presented in the subsequent nodes. In addition to the ease of interpretation, the main advantage of a decision tree is that decision is made based on the most important attribute. The decision tree presents the attributes in order of importance and allows for knowing the factors of greater influence in the study (Lemos, 2003).

Decision trees are structures that can be used for training people to learn from the generated information and for decision-making. The learning process occurs by observing the world interactions and from the internal process of decision-making. The decision tree uses a “divide and conquer” strategy as a complex problem is decomposed into simpler sub-problems with the same technique applied to each sub-problem. The discrimination power of a decision tree come from the space defined by the attributes divided into subspaces, and from the association of one class to each space (Gama, 2000).

According to Garcia (2003) the decision trees are induced from a set of training examples whose classes are previously known and made of:

(i) Nodes representing the predictors;
(ii) Branches starting from the nodes and receiving the possible values for these predictors;
(iii) Tree leaves representing the different classes of a training set, that is, each leaf is associated with one class.

Each path of the tree, from the root to the leaf, means a rule of classification. In the decision tree, each node should be associated to one attribute that must be the most predictive among those attributes not considered yet on the path (Lemos, 2003).

According to Carvalho (2005), the classification of an example occurs when the example “goes through” the tree from the root node, travelling the branches connecting the nodes following the conditions of these branches. When reaching one leaf node, the class that labels the leaf is attributed to this example.

The function of the decision tree is to create subsets from a training set containing examples of unique class to build a model for further classifications (Quinlan, 2014). To build a decision tree, the fundamental idea is to choose an attribute; develop the tree by adding a branch to each attribute value; make the example “goes through” each leaf; when all examples reach the same class associate the class with the leaf, otherwise, repeat all these four steps again. Each tree path, from the root to the leaf, corresponds to a rule of classification.

According to Pereira et al. (2007), the induction of rules corresponds to the process of discovering patterns among data, that is, finding rules of prediction as *if…then,* where *“if”* is the specific condition of some attribute and *“then”* is the action of the rule that predicts a value for a given attribute.

Lemos et al. (2005) affirm that decision trees are commonly substituted by rules because, as demonstrated in some applications, the trees tend to grow much, and rules can be easily modulated. Building a decision tree should start from a training set containing examples that are previously known and historic data, both negative and positive. To complete a decision tree with a high prediction power, it is necessary to choose properly the attributes that will be used for training groups so that these tests generate a tree with the smallest possible number of subsets, where each leaf has the greatest possible significant number of examples. Ideally, the tree should be as small as possible. Because analyzing all possibilities is impracticable, several methods were developed and applied to select attributes and types of tests. At the same time, they have agreed in two points: the division in which are kept all the proportions of classes in all partitions is useless; and the division in which the examples in each partition are of the same class is of greatest importance. Once the choice is made, the other possibilities are not explored further (Lemos, 2003). Before selecting an attribute, it is necessary to know two concepts: entropy and information gain. Entropy is the measure of randomness of one variable. It is also defined as the indicator of the homogeneity of the examples from a dataset. It indicates the purity or impurity of examples (Osório, 2000).

The construction of a decision tree is guided by the goal of reducing the entropy (randomness), that is, the difficulty to predict object variables. To reduce entropy, information is gained by the partition of examples according to the attribute values (Carvalho, 2005). The information gain represents the difference between the amount of information necessary for correct prediction and the corresponding accumulated amounts of segments resulting from a new test to determine the value of a given attribute.

Some advantages of decision trees is that: it is a non-parametric method and then no particular distribution for data is assumed so that models can be built for any function as the number of training examples is efficient; the structure of the decision tree does not depend on the scales of variables; it allows a high level of interpretability; decision trees are efficient to build models and robust with respect to extreme points and redundant or irrelevant attributes.

When studies about insolvency of credit unions are analyzed, four cases are worth mentioning: The first by Bressan et al. (2004b) evaluated the financial health and economic situation of rural credit cooperatives of Minas Gerais State (Brazil); in this work, a logistic regression was applied to the sample of rural credit cooperatives in the period of 1998-2001. The second study, also by Bressan et al. (2004a), evaluated the insolvency of rural credit cooperatives integrating the Crediminas system by using the Cox proportional hazards model. The third study by Bressan (2009) computed the possibility of insolvency of credit unions integrated to the SICOOB-Brasil and SICCOB-Crediminas systems by using indicators from the PEARLS system and Logit Model. The fourth study by Gozer et al. (2014) developed and compared statistical models using the techniques of Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) to investigate which one offers the best results in evaluating insolvency of mutual credit unions.

With respect to models for predicting insolvency, there are several approaches for either capital firms or credit unions, in which diverse financial indicators have been used as the indicators of the PEARLS system were selected since they are recommended by the World Council of Credit Unions (WOCCU) (Gozer et al., 2014; Bressan et al., 2015).

Araújo (2011) mentioned that the Federal Financial Institutions Examination Council (FFIEC) proposed the

Uniform Financial Institutions Rating System (UFIRS), also known as CAMELS. The acronym CAMELS comprises a set of financial indicators used for monitoring U.S. financial institutions.

The CAMELS is not applied to evaluate credit unions, but WOCCU, an international agency created to promote credit unions, has developed an adaptation for these cases. Then, the PEARLS system for financial analysis of credit unions is widely used by its affiliates (Bressan, 2009; Bressan et al., 2015).

According to Bressan (2009), the PEARLS system is an acronym for a group of indicators deriving from the evaluation of some key operational areas of credit unions: Protection, effective financial structure, assets quality, rates of return and costs, and signs of growth. Monitor the performance of the credit union is the main objective of the PEARLS system (Gozer et al., 2014). It was designed to be a tool that goes beyond the simple identification of the problem and to help managers finding meaningful solutions to institutional problems. The PEARLS system can identify whether a credit union is based on a financial imbalance and the probable causes of this problem. This system allows managers to quickly and accurately identify problematic areas and make the necessary corrections before the problem becoming serious. Therefore, the PEARLS system is a tool that can precede problems, generating extremely useful information for the financial management of credit unions (Gozer et al., 2014; Bressan et al., 2015). The PEARLS system also aims to standardize financial indicators, formulas and criteria to evaluate the operations of credit unions and build for this, a global financial language (Richardson, 2002).

The PEARLS system is adopted by about 97 countries in Africa, Asia, Caribe, Europe, North America, Latin America and Oceania. However, it is not used in Brazil (Bressan et al., 2011). Bressan et al. (2011), following the recommendations of Vasconcelos (2006) and the foundations of Bressan (2002), Richardson (2002) and Ribeiro (2008), created 39 financial indicators within the PEARLS classification to evaluate the Brazilian credit unions. These indicators have enabled comparisons and financial analysis among credit unions in Brazil and abroad. The indicators were grouped into key-operating areas of the credit unions and addressed to the *Plano Contábil das Instituições Financeiras do Sistema Financeiro Nacional – COSIF* (Accounting Plan of National Institutions of the Financial System).

METHODS

The evaluation of insolvency of credit unions sampled in this study uses the financial indicators of the PEARLS system (Gozer et al., 2014). This system was created in the late 80´s by the WOCCU to be used with credit unions. PEARLS is an acronym for a group of indicators deriving from the evaluation of some key operational areas of single credit unions: Protection, effective financial structure, assets quality, rates of return and costs, and signs of growth. From those 39 indicators proposed by the PEARLS system, 27 were selected due to the availability of information on the financial statements prepared and provided by the BACEN website (www.bacen.gov.org).

The WEKA machine learning software (Waikato Environment for Knowledge Analysis) was used for data mining. The WEKA is formed by a set of algorithms of several techniques to solve real problems of data mining. Its development occurred in the academic environment, at the University of Waikato in New Zealand by 1999.

Almeida (2010) evidences that this software began to be described by 1993 and after it were acquired by a company in 2006. The WEKA is licensed under the General Public License, so it is possible to study it and modify its respective source code. Still, according to the author, WEKA aims to aggregate algorithms from different approaches in the subarea of artificial intelligence dedicated to the study of machine learning. This subarea aims to develop algorithms and techniques that allow a computer to "learn" in the sense of gaining new knowledge, either inductively or deductively.

According to Almeida (2010), the WEKA contains tools for preprocessing, classification, regression, clustering, association rules and data view. It is also suitable for developing machine learning systems.

In decision tree technique a characteristic is chosen and a particular value of the characteristic is chosen to partition the cases into two subsets. The characteristic becomes a decision node and each decision, indicated by a particular value of the characteristic, forms a branch. Each branch leads to a different characteristic. Again a particular value of this characteristic is chosen to partition the subsets into further subsets and so on (Kotsiantis, 2013; Islam and Habib, 2015).

Three decision trees were built with the J48, ADTree and LADtree algorithms. The decision trees algorithms details are found in Barros et al. (2012) and Kotsiantis (2013). For its construction, it was used a paired data base of 62 credit unions, from which 31 are solvent and 31 insolvent. Cooperatives were considered insolvent when stopped sending financial statements to the Central Bank during a period of at least 10 years, thus characterizing a situation of operational disruption. Kappa statistics was used to support the selection of the most efficient decision tree. Kappa statistics evaluates the level of agreement of a classification task, in which, by means of different techniques, considers only the concordance among classifiers, indicating that the classified data have certain cohesion.

RESULTS AND DISCUSSION

Credit unions are very importance in the international financial scene, and have demonstrated its growth potential in Brazil (Bressan et al., 2015). The importance of the expansion of credit unions is, above all, the possibility of extending access to finance today for those that are not served by the traditional banking sector, contributing to the process of financial disintermediation (Bressan et al., 2015). Therefore, the analysis of the financial structure of the central cooperatives, which are responsible for assisting the management of individual cooperatives, providing input to policies and interventions by the Central Bank, and helping the financial manager to monitor the status of their institutions, allowing also greater security to economic agents operating with these institutions (Bressan et al., 2015).

Due the importance of credit unions, we evaluation of

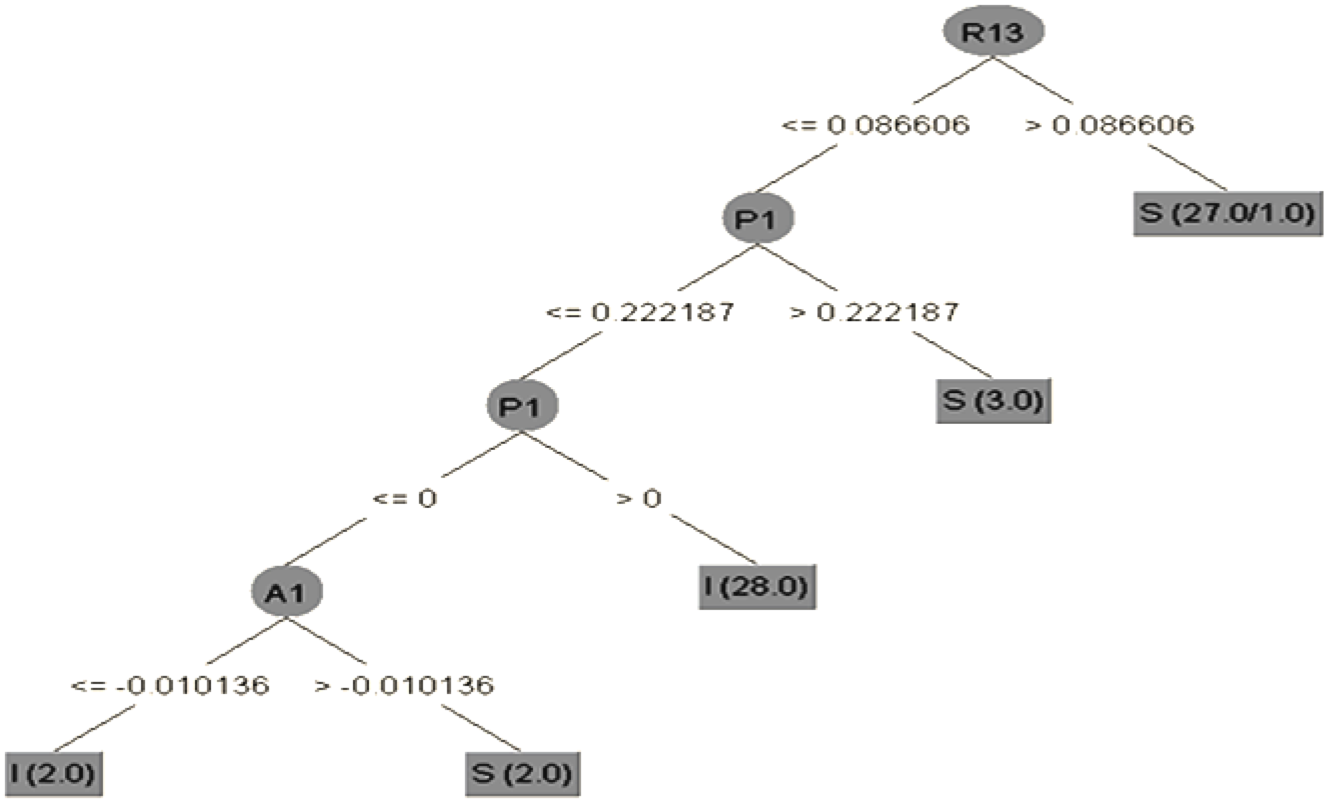

Figure 1. Decision tree constructed with the J48 algorithm of mutual credit unions in the Paraná State, Brazil.

Table 1. Kappa coefficient.

Kappa coefficient	Level of accuracy of the classifier
< 0.20	Weak
0.21 - 0.40	Regular
0.41 - 0.60	Moderate
0.61 - 0.80	Good
< 0.81	Excellent

Source: Oliveira (2004).

the insolvency of mutual credit unions in the Paraná State (Brazil) by application of the data mining using decision trees approach. The information required to build the models were obtained from indicators applied to a sample of 62 mutual credit unions from which 31 are solvent and 31 are insolvent. For analysis and discussion of the results it is presented in sequence the construction of the three decision trees. First, was constructed with the J48 algorithm (Figure 1); second with the ADTree (Figure 2) and third, with the LADtree algorithm (Figure 3).

The Kappa statistics, which defines the level of accuracy of the classifier, was applied to help the selection of the most efficient decision tree. According to this statistics, the decision tree constructed with the J48 algorithm obtained an indicator value of 0.6182, that one constructed with the LADtree algorithm achieved a value equal to 0.5249, while the decision tree built with the ADtree algorithm, a value of 0.8108.

Given these results and according to the level of accuracy of the classifier presented, the decision tree constructed with the J48 algorithm achieved a classification level considered good (Figure 1 and Table 1). The decision tree built with the ADTree algorithm achieved a level of classification considered excellent (Figure 2 and Table 1). The decision tree built with the LADtree algorithm achieved a classification level considered moderate (Figure 3 and Table 1). Therefore, the decision tree built with the ADTree algorithm was

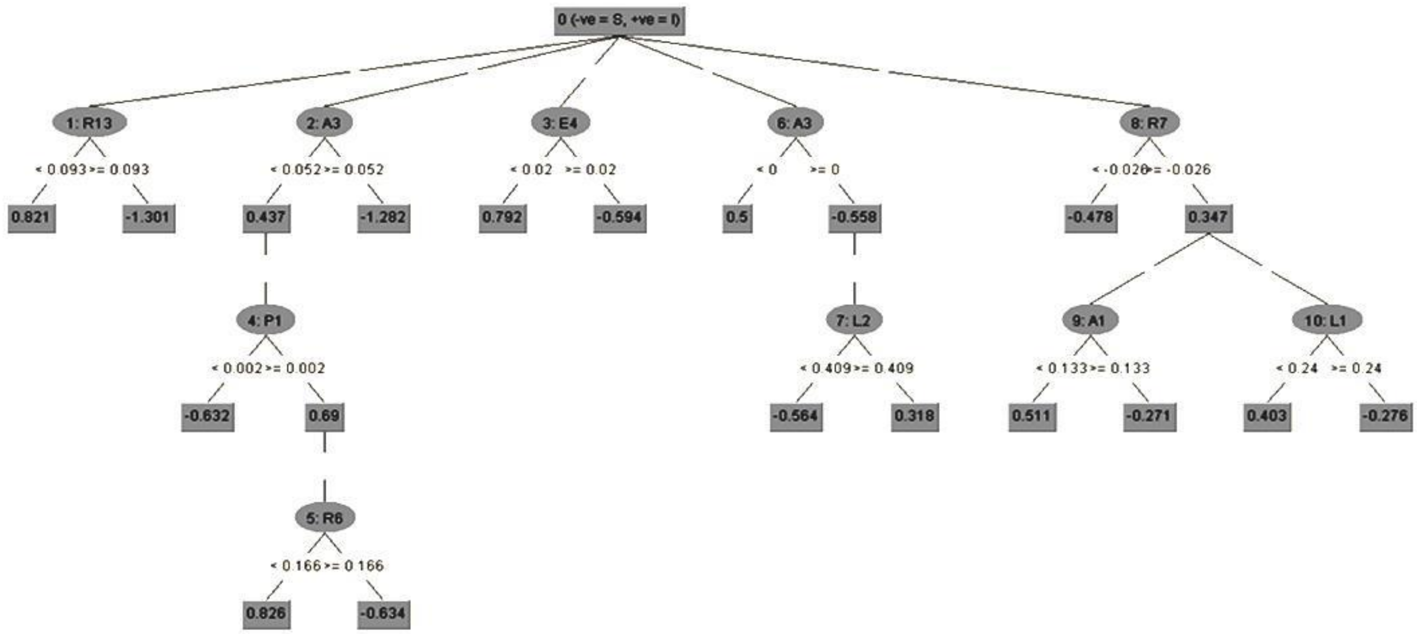

Figure 2. Decision tree constructed with the ADTree algorithm of mutual credit unions in the Paraná State, Brazil.

chosen for the best performance.

The decision tree built with the ADTree algorithm presenting the best level of accuracy of the classifier indicated by the Kappa coefficient is not associated with a class of an example as the value of a leaf, but with the signal obtained by summing all prediction values of the nodes traversing from the root until one leaf of the tree, therefore, the training started from the zero point (Figure 2). Negative values were addressed to the insolvent cooperative and positive values to the solvents. Therewith, it can be observed in the first node that if the R13 indicator (Administrative Costs / Average Total Assets) is lower than 0.093, the cooperative is considered solvent within a confidence level of 0.821; but if this indicator exceeds 0.093, the cooperative is insolvent within a confidence level of 1.301 (Figure 2).

The A3 indicator (Net Institutional and Transitory Capital + Non-Interest-bearing Liabilities / Non-earning Assets) appears in the second node of the tree (Figure 2). If it is higher than 0.052, the cooperative is solvent within a confidence level of 1.282; but if it is lower than 0.052 and the P1 indicator (Allowance for Loan Losses/Delinquency > 12 months) is lower than 0.002, the cooperative is insolvent within a confidence level of 0.195. Still in the same node, if the A3 indicator is higher than 0.052, the P1 higher than 0.002 and the R6 indicator (Total Interest Cost on External Credit / Average External Credit) higher than 0.69, the cooperative is solvent within a confidence level of 1.953; and if the R6 is lower than 0.166, the cooperative is solvent within a confidence level of 0.493.

Observing the third node of the decision tree, if the R4 indicator (Total Non-financial Investment Income / Avg. Non-financial Investments) is higher than 0.02, the cooperative is solvent within a confidence level of 0.792; and if it is lower than 0.02, the cooperative is insolvent within a confidence level of 0.594 (Figure 2).

In the fourth node of the tree, if the A3 indicator is below zero the cooperative is insolvent within a confidence level of 0.5, and if it is above zero but the L2 indicator (Liquidity Reserve / Savings Deposits) is lower than 0.409, the cooperative is insolvent within a confidence level of 1.122; when the L2 indicator is higher than 0.409, the cooperative is insolvent within a confidence level of 0.240 (Figure 2).

The R7 indicator (Total Interest (Dividend) Cost on Shares / Average Member Shares) appears in the end node of the tree (Figure 2). When it is lower than 0.026, the cooperative is insolvent with a confidence level of 0.478. In the same node, if the R7 is higher than 0.026 and the A1 (Total Loan Delinquency / Gross Loan Portfolio) is higher than 0.133, the cooperative is solvent within a confidence level of 0.858. On the same way, if the A1 is lower than 0.133, the cooperative is solvent within a confidence level of 0.076. In the same node, if the R7 is higher than 0.026 and the L1 (ST Investments + Liquid Assets – ST Payables/Savings Deposits) is higher than 0.24, the cooperative is solvent within a confidence level of 0.750; and if it is lower than 0.24, the cooperative is solvent within a confidence level of 0.071. Observing the results, the indicators related to returns and costs (R13, R6 and R7) were the most present on the evaluation of insolvency of a credit union.

It is worth noting the importance of the R13 indicator

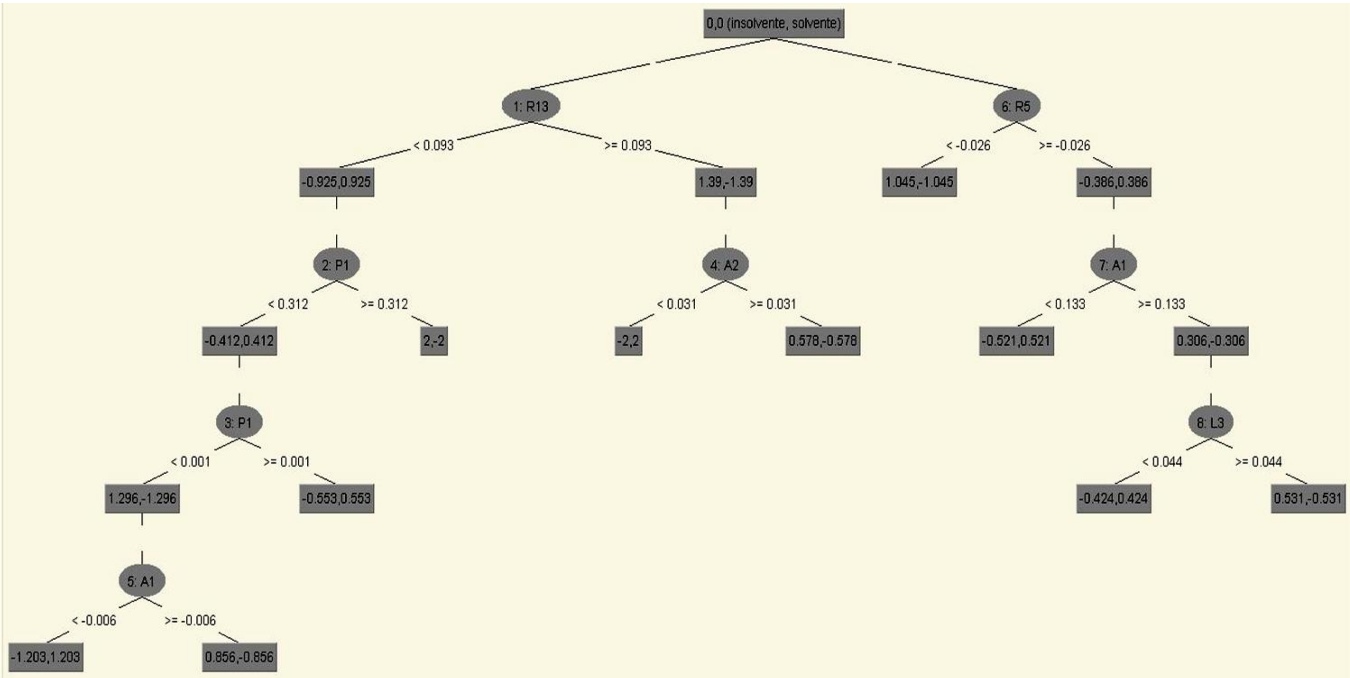

Figure 3. Decision tree built with the LADtree algorithm of mutual credit unions in the State of Paraná.

that appeared alone in one node of the tree. The model evidences that the control of the administrative costs is important to assess the insolvency of the cooperatives in this study since this indicator can identify the management efficiency of cooperatives and consider it as a relevant factor for evaluating their state of insolvency.

From these results, it is possible to affirm that the indicators related to quality of assets are also important for the accuracy in predicting the insolvency state of cooperatives as the A1 and A2 indicators were present in two nodes of the tree. In addition, the liquidity flow indicators, L1 and L2, were also present. With respect to indicators of protection, only P1 appeared, and no indicator of growth was present in the tree.

Finally, the path with the highest confidence level (1.953) to evaluate insolvency of credit unions was that of the second node of the decision tree: The A3 higher than 0.052, P1 higher than 0.002 and R6 higher than 0.166 (Figure 2).

Bressan (2009) examined the insolvency of credit unions integrated to the SICOOB-Brasil and SICOOB-Crediminas systems using the Logit Model and identified the following indicators of the PEARLS system as determinant to predict insolvency: P2 (Net Allowance for Loan Losses/Delinquency of 1-12 months), E4 (Non-financial Investments/Total Assets), A3 (Assets not intended to target activity/ Cooperative's total assets) and R11 (Non-recurring Income or Expense/Average Total Assets).

The indicators pointed in the present study were determinant for predicting insolvency of credit unions were confirmed in the present study although it is using a different technique – the decision tree. In this sense, this study highlights the following indicators: A3 (Net Institutional and Transitory Capital +Non Interest-bearing Liabilities/ Non-earning Assets), E4 (Non-financial Investments/Total Assets), R13 (Administrative Costs / Average Total Assets) and P1 (Allowance for loan losses / Delinquency > 12 months). The indicator P2 (Net Allowance for loan losses / Delinquency of 1-12 months) was not present in the decision tree because the financial statements provided by BACEN did not allow its calculation.

Final considerations

In this study was assumed the possibility of developing a model for evaluating the state of insolvency of mutual credit unions by using the decision tree technique. Trying to study efficiency in the cooperative credit sector has led to adopt new technology and managerial knowhow. Among the tools that facilitate efficiency, data mining has stood out in recent years as a sophisticated methodology to search for knowledge that is “hidden” in organizations' databases (Sousa and Figueiredo, 2014). In our knowledge, the present study is the first to evaluate insolvency of mutual credit unions by application of the data mining using decision trees approach.

Others studies was also applied data mining using approach and used decision trees, but for others propose e.g. Islam and Habib (2015) used data mining approach to predict prospective business sectors for lending in retail banking using decision tree attempting to build up a model to predict prospective business sectors in retail banking. Decision trees were built from the J48, ADtree and LAD tree algorithms. The tree built with the ADtree algorithm had the best performance according to the Kappa statistics, showing an excellent level of accuracy for the classifier.

The decision tree of best performance, built with the ADtree algorithm, had the highest predicting power for the following indicators: R13 (Administrative Costs / Average Total Assets), R6 (Total Interest Cost on External Credit/Average External Credit), R7 (Total Interest (Dividend) Cost on Shares/Average Member Shares), A1 (Total Loan Delinquency/Gross Loan Portfolio), A3 (Net Institutional & Transitory Capital + Non-Interest-bearing Liabilities / Non-earning Assets), L1 (ST Investments + Liquid Assets – ST Payables / Savings Deposits), L2 (Liquidity Reserve / Savings Deposits) and P1 (Allowance for Loan Losses / Delinquency > 12 months).

The path with the highest confidence level (1.953) to evaluate insolvency of credit unions was that of the second node of the decision tree: The A3 (Net Institutional and Transitory Capital + Non Interest-bearing Liabilities/ Non-earning Assets) higher than 0.052; P1 (Allowance for Loan Losses/Delinquency > 12 months) higher than 0.002; and R6 (Total Interest Cost on External Credit/Average External Credit) higher than 0.166. From the literature survey made on this topic, it was noticed that there is much heterogeneity among studies due to the use of different concepts of insolvency, sample composition, techniques (discriminant analysis, conditional probability, artificial neural networks, Cox model), and number and type of selected variables (traditional indicators of financial statements, the PEARLS system) (Gozer et al., 2014).

Many techniques are available for research, all presenting weak and strong points, thus, there is no agreement on what is the best. The selection of independent variables in certain studies has been made with the help of econometrics. Such selection has been based on the performance of variables tested on previous studies or, as mentioned by some authors, on the availability of information. This study has implications for developing a general theory of corporate insolvency as well as for further researches considering the option of selecting and deciding on aspects previously mentioned or on information available for research.

Conflict of Interest

The authors have not declared any conflict of interest.

ACKNOWLEDGEMENT

The authors thank the Universidade Paranaense - UNIPAR by supporting this work.

REFERENCES

Almeida ED (2010). Algoritmos de classificação com a opção de rejeição. Universidade do Porto, Portugal. MSc Dissertation. 19pp.

Araújo MBV (2011). Informações contábeis e o risco de insolvência de cooperativas de crédito. Universidade de São Paulo, São Paulo, Brazil. MSc Dissertation. 143pp.

Barros RC, Basgalupp MP, de Carvalho ACPLF, Freitas AA (2012). A Survey of Evolutionary Algorithms for Decision-Tree Induction. IEEE transactions on systems, man and cybernetics. Part C, Appl. Rev. 42:291-312.

Braga MJ, Bressan VGF, Colosimo EA, Bressan AA (2006). Investigating the solvency of Brazilian credit unions using a proportional hazard model. Ann. Public Coop. Econ. 77:83-106.

Bressan VGF (2002). Análise de insolvência das cooperativas de crédito rural do Estado de Minas Gerais. Universidade Federal de Viçosa, Minas Gerais, Brazil. MSc Dissertation. 123pp.

Bressan VGF (2009). Seguro depósito e moral Hazard nas cooperativas de crédito brasileiras. Universidade Federal de Viçosa, Minas Gerais, Brazil. Doctor Thesis. 371pp.

Bressan VGF, Braga MJ, Bressan AA (2004b). Análise do risco de insolvência pelo modelo de Cox: uma aplicação prática. Rer. Adm. Empres. 44:83-96. http://rae.fgv.br/sites/rae.fgv.br/files/artigos/10.1590_S0034-5902004000000008.pdf

Bressan VGF, Braga MJ, Bressan, AA, Resende Filho MA (2011). Uma aplicação do sistema PEARLS às cooperativas de crédito brasileiras. R. Adm. 46:256-274.

Bressan VGF, Braga MJ, Lima JE (2004a). Análise de insolvência das cooperativas de crédito rural do Estado de Minas Gerais. Est. Econ. 34:553-585. http://www.scielo.br/pdf/ee/v34n3/v34n3a06.pdf

Bressan VGF, Bressan AA, Oliveira PHM, Braga MJ (2015). Quais indicadores contábeis financeiros do sistema pearls são relevantes para análise de insolvência das cooperativas centrais de crédito no Brasil? Contab Vista & Rer. 25:74-98.

Carvalho DR (2005). Árvore de decisão/algoritmo genético para tratar o problema de pequenos disjuntos em classificação de dados. Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil. Doctor Thesis. 162pp.

Chaves SS (2009). Cooperativismo de crédito e empresas de pequeno porte em arranjos produtivos locais. Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil. Doctor Thesis. 236pp.

Gama J (2000). Árvores de decisão. Retrieved from: <http:// www.liacc.up.pt/~jgama>. Accessed on 14 Ago. 2014.

Garcia SC (2003). O uso de árvores de decisão na descoberta de conhecimento na área da saúde. Universidade Federal do Rio Grande do Sul, Porto Alegre. Brazil. MSc Dissertation. 87pp.

Gozer IC, Albuquerque ARP, Isotani S, Gimenes RMT, Moreira WH, Alberton O, Menezes EA (2014). Evaluation of insolvency in mutual credit unions by the models of artificial neural networks and support vector machines. Afr. J. Agric. Res. 9:1227-1237.

Han J, Kamber M (2006). Data mining: concepts and techniques. 2.ed. University of Illinois at Urbana-Champaing; Elsevier, New York. 703 pp.

Islam MR, Habib MA (2015). A data mining approach to predict prospective business sectors for lending in retail banking using decision tree. Int. J.Data Mining Knowledge Manag. Process. 5:13-22.

Kotsiantis SB (2013). Decision trees: a recent overview. Artif. Intell. Rev. 39:261-283.

Lemos EP (2003). Análise de crédito bancário com uso de data mining: Redes Neurais e Árvores de Decisão. Universidade Federal do Paraná, Curitiba, Brazil. MSc Dissertation. 147pp.

Lemos EP, Steiner MTA, Nievola JC (2005). Análise de crédito bancário por meio de redes neurais e árvores de decisão: uma aplicação simples de *data mining*. Revista de Administração. 40:225-234.

Lima MGF (2007). Árvore de decisão aplicada a bancos de dados de incubatório de matrizes de postura. Universidade Estadual de Campinas, Campinas, Brazil. MSc Dissertation. 80pp.

Oliveira AJF (2004). Método para avaliação de risco operacional em bancos. Universidade Federal de Santa Catarina, Florianópolis, Brazil. MSc Dissertation. 143pp.

Osório F (2010). Sistemas adaptativos inteligentes – indução de árvores de decisão. Retrieved from: http://www.inf.unisinos.br/~osorio/sadi.html>. Accessed on 15 Oct. 2014.

Pereira JM, Domínguez, MÁCZ, Ocejo JLS (2007). Modelos de previsão do fracasso empresarial: aspectos a considerar. Tékhne - Revista de Estudos Politécnicos. 4:111-148.

Quinlan JC (2014). C4.5: programs for machine learning. Elsevier.

Ribeiro DM (2008). Insolvência de cooperativas de crédito: uma aplicação do modelo Cox com covariâncias dependentes do tempo. Universidade Federal de Minas Gerais, Belo Horizonte, Brazil. Monograph. 94pp.

Richardson DC (2002). PEARLS Monitoring System. Word Council of Credit Unions. Toolkit series number 4. Retrieved from: <http://www.woccu.org/bestpractices/pearls/pearlsmonograph>. Accessed on 20 Oct. 2014.

Sousa MM, Figueiredo RS (2014) Credit analysis using data mining: application in the case of a credit union. J. Info. Sys. Technol. Manage. 11:379-396.

Vasconcelos RWB (2006). Identificação de indicadores econômico-financeiros para análise de cooperativas de crédito singulares ou centrais. Belo Horizonte: Departamento de Supervisão Indireta e Gestão da Informação (DESIG), Banco Central do Brasil.