academicJournals

Vol. 9(16), pp. 1227-1237, 17 April, 2014
DOI: 10.5897/AJAR2014. 8654
Article Number: 7858DC443953
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 the models of artificial neural networks and support vector machines

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]Department of Biotechnology Applied to Agriculture; Paranaense University – UNIPAR, Umuarama, Paraná, Brazil.
[6]Universidade Federal de Santa Catarina, Santa Catarina, Brazil.

Received 1 March, 2014; Accepted 7 April, 2014

This study aimed to develop and compare 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. The information required to build the models were obtained with a sample of 62 credit unions (31 solvent and 31 insolvent) to which financial indicators of the PEARLS (Protection, Effective financial structure, Asset quality, Rates of return and cost, Liquidity and Signs of growth) system were calculated. The RBF network, multilayer perceptron, multilayer perceptronCS and LibSVM algorithms were used to obtain the ANNs and SVMs; for each algorithm, the ANNs were built with three groups of indicators (27, 11 and 10 indicators). This is the first study done with ANNs in Brazilian credit unions. When analyzing the results of ANNs and SVMs, the superiority of the SVMs as binary classifier for evaluating insolvency was evidenced, since its LibSVM algorithm showed the best results in all assessments of performance proposed in this study. The only LibSVM indicator with performance inferior to ANNs was the error rate of the negative class which indicates those negative class data that were classified incorrectly.

Key words: Insolvency, credit unions, artificial neural networks, support vector machines.

INTRODUCTION

Cooperatives are currently an alternative means to make credit accessible and include people who are at the boundaries of the National Financial System in the financial market; especially to benefit those small urban

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

and rural entrepreneurs by raising funds at lower interest rates than the average rates charged on the market.

According to Bressan et al. (2011), one of the biggest challenges of the credit union is to build management models adjusted to the particularities of the business segment meeting all requirements and doctrinal principles of the regulatory agency, in this case of Central Bank of Brazil (BACEN).

Studies making use of well-designed statistics to predict insolvency in cooperatives have been performed long ago. Initially, univariate techniques were applied in this type of study, and then, multivariate analyzes appeared bringing more suitable solutions. Currently, techniques that make use of artificial intelligence and need a great capacity for computational processing have been largely adopted (Azmathullah et al., 2005; 2006; 2008; Hsiao and Whang, 2009; Jardin, 2010; Azmathullah and Wu, 2011; Karra and Krichene, 2012).

By observing this progress, we see continuous improvements in the results achieved on this theme. Today, it is possible to affirm that the financial statements of the cooperatives when properly analyzed are sources of information for predicting insolvency more accurately (Hsiao and Whang, 2009).

Quantitative models for insolvency prediction may be built from financial indicators obtained from the financial statements of a sample of solvent or insolvent cooperatives. Statistical models work to distinguish the characteristics of solvent and insolvent cooperatives. The final result of this processing is a mathematical formula that predicts the future situation of a cooperative within a certain margin of error.

This study aimed to evaluate the state of insolvency of the mutual credit unions in the Paraná State (Brazil) by the methods of artificial neural networks (ANNs) and support vector machines (SVMs). This is the first study done with ANNs in Brazilian credit unions.

Theoretical backgrounds

Bressan et al. (2004a) consider that any method to evaluate the risks of bankruptcy or insolvency of a company should meet the following stages: i) obtaining a sample containing solvent and insolvent companies; ii) choosing variables from financial statements to indicate company insolvency which will be called predictor variables; iii) choosing a mathematical or statistical model to fit the variables obtained from financial statements; and iv) validating the model in order to check the model ability to discriminate categorical variables.

In Brazil, most studies on models for insolvency prediction have proposed to analyze the model variables. Normally, it is done based on a statistical technique where the researchers aim to find a set of appropriate financial indicators to predict accurately the company's financial health in a given period in the future (Wu et al., 2007; Fernandez et al., 2013). The statistical modeling commonly used with this purpose is the discriminant analysis. On the same line, the insolvency models have showed increasingly relevance in financial analysis mainly by taking into account the credit and life insures risks (Wu et al., 2007; Hsiao and Whang, 2009; Jardin, 2010; Ribeiro et al., 2012).

With respect to credit unions, Table 1 shows logistic regression analysis performed in four works, the Cox proportional-hazard regressions applied in three works, and discrimant analysis in one work (Araújo, 2011). Note that no study with ANNs was found within the analyzed period for Brazilian credit unions.

With respect to studies on insolvency of credit unions, three studies are worth mentioning. The first by Bressan et al. (2004a) describes an economic-financial evaluation of rural credit cooperatives in the Minas Gerais State (Brazil) made by logistic regression on a 1998 to 2001 sample of the rural credit cooperatives of Minas Gerais. The second, also by Bressan et al. (2004b), evaluates the insolvency of rural credit cooperatives of the Crediminas system by using the Cox proportional-hazard model. The third work, developed by Bressan (2009; 2011), uses PEARLS (Protection; Effective Financial Structure; Assets Quality; Rates of Return and Costs; Liquidity; and Signs of Growth) indicators with Logit Model to predict the possibility of insolvency of credit unions that participate in the SICOOB-Brasil and SICCOB-Crediminas systems.

According to Bressan (2009, 2011), the PEARLS system is one acronym of a group of indicators used by the Word Council of Credit Unions (WOCCU) (2002) since 1990, which originates from the following key-operating areas of the credit union.

Monitoring the credit union performance is the main objective of the PEARLS system. It was designed to go beyond the identification issue, helping managers to find significant solutions to solve institutional problems. The PEARLS system can identify if a credit union is based on weak capital and indicate the causes of such weakness. This system allows managers to identify problems and solutions quickly and accurately, and the best actions to be taken before the problems becoming serious.

Another objective of the PEARLS system is to standardize financial indicators and formulas to eliminate criteria used by local credit unions to evaluate operations. The system also creates a universal financial language to evaluate credit unions worldwide – with an easy-access language that can improve the communication and information uniformity (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 to analyze credit unions in Brazil (Bressan et al., 2011).

Bressan et al. (2011) have created 39 financial indicators from the PEARLS system to evaluate the Brazilian credit unions. These indicators have enabled

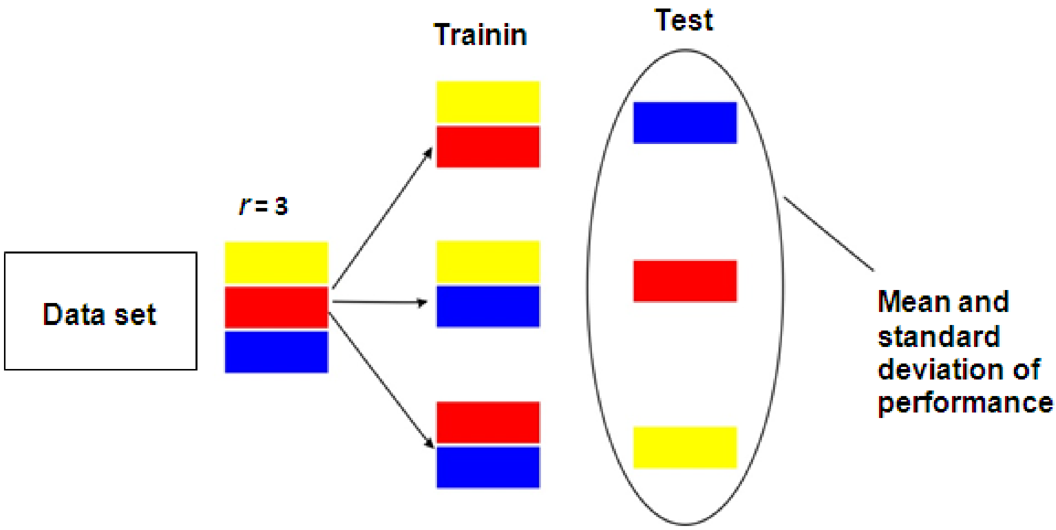

Figure 1. Cross Validation Method. Source: Carvalho (2011).

comparisons and financial analysis among credit unions in Brazil and abroad (Karaa and Krichene, 2012).

The indicators were grouped into key-operating areas of the credit unions and linked in the *Plano Contábil das Instituições Financeiras do Sistema Financeiro Nacional – COSIF* (Accounting Plan of National Institutions of the Financial System) (Bressan et al., 2010).

METHODS

Data gathering was made based on information from financial statements of a sample of 62 credit unions from which 31 are solvent and the other 31 are insolvent. The cooperatives that did not send their financial statements to BACEN in a period of at least 10 years were considered insolvent since it characterizes an operational disruption.

According to Araújo (2011), this concept suggests that interrupting the sending of financial statements implies lack of interest of the cooperative managers in being accountable to the regulator, and this usually occurs before the formal closure of the cooperative – a typical profile of insolvency.

The evaluation of insolvency of cooperatives was based on the financial indicators of the PERALS system, as mentioned in the theoretical background of this study. Twenty-seven indicators were selected from 39 proposed by the PEARLS system according to the availability of information on the financial statements prepared and provided by the BACEN website.

To obtain the best model of data classification by artificial neural networks (ANNs) and support vector machines (SVMs), the re-sampling technique by the cross validation method was used (Figure 1) (Carvalho, 2011) together with the evaluation of the generated models by visualizing points in two-dimensional space (*x, y*) on receiver operating characteristics (ROC) curve (Figure 2) (Kohonen, 1988).

In the *r - fold cross-validation* method, the data set is partitioned into *r* subsets of similar sizes. The objects of *r-1* partitions are used for training a predictor variable which is tested in the remaining partitions. This process is repeated for *r* times, each cycle with a different partition to be tested. The final performance is taken from the averaged performance observed in each subset of test (Carvalho, 2011).

The re-sampling is indicated when the sample is not representing the population well and then an estimate closer to the study population is needed. The *k - fold cross-validation* consists of randomly selecting a group of data extracted from the dataset to train and test predictors. The iteration was made 10 folds changing the number of test samples and tests, but keeping the same training dataset (Han and Kamber, 2006).

In addition to compute algorithms by the ANNs and SVMs, a ROC curve was also designed. According to Han and Kamber (2006), it is important to validate the results in order to quantify the discriminating power for prediction and identify the accuracy of a method or procedure in certain analysis. However, it should be considered that only a single quantification of "misses and hits" in a given test group is not necessarily reflecting the efficiency of the process. The quantification is also highly depending on the quality of data distribution through the test group. On the ROC curve, a perfect classifier is showed by a horizontal line on the top of the graph. However, it would be hardly achieved. ROC curves are considered good when positioned between the diagonal and the perfect lines, that is, the farther from the diagonal line, the better the algorithm.

According to Carvalho (2011), a classification conflict in the two classes, one positive (+) and the other negative (-), generates a confusion matrix as shown in Table 2, where: TP = the number of true positives, which is the positive data correctly classified; TN = the number of true negatives, which is the negative data correctly classified; FP = the number of false positives, data whose the class is negative but were incorrectly classified as positive; and FN = the number of false negatives, data whose the class is positive but were incorrectly classified as negative.

According to Monard and Baranauskas (2003), it is possible to calculate other performance measures from the confusion matrix:

a) Error rate in the positive class: The proportion of positive class data that were incorrectly classified by the predictor $\hat{f}$ also known as false negative rate (FNR):

$$(\hat{f}) = \frac{\mathrm{FN}}{\mathrm{TP}+\mathrm{TN}}$$

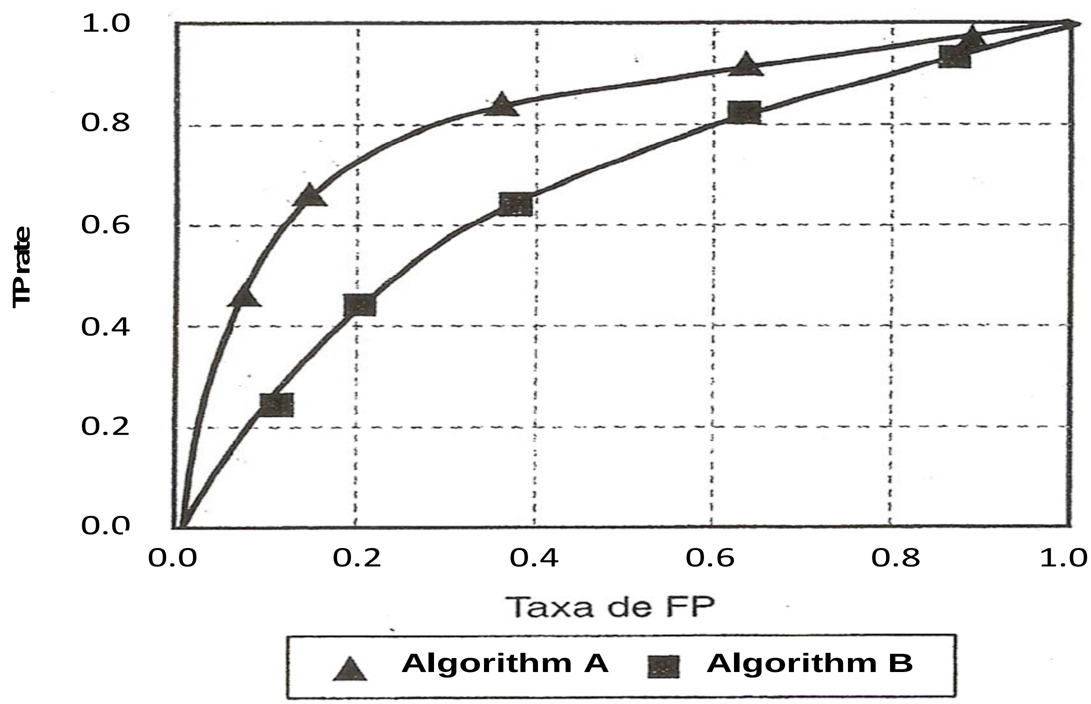

Figure 2. Example of a receiver operating characteristics (ROC) curve. Source: Carvalho (2011).

Table 1. Studies on credit union in solvency in Brazil.

Method	References
Logistic regression and cox proportional hazard	Bressan (2002)
Discriminant analysis	Pinheiro (2003)
Logistic regression and cox proportional hazard	Bressan et al. (2004a)
Cox proportional hazard	Bressan et al. (2004b)
Cox proportional hazard	Ribeiro (2008)
Logit model	Bressan (2009)
Logistic regression	Ferreira (2010)

Source: Araujo (2011).

b) Error rate in the err_{+} negative class: Proportion of negative class data incorrectly classified by $\hat{f}$, also known as false positive rate (FPR): $err_{-}\left(\hat{f}\right) = \frac{\mathrm{FP}}{\mathrm{FP+TN}}$,

c) Total error rate: Given by the sum of the values of the secondary diagonal of the matrix divided by the sum of values of all components of the matrix: $err\left(\hat{f}\right) = \frac{\mathrm{FP+FN}}{n}$,

d) Precision rate or overall precision: Obtained by adding the values of the main diagonal of the matrix divided by the sum of values of all components of the matrix: $prec_{-}\left(\hat{f}\right) = \frac{\mathrm{TP+TN}}{n}$,

e) Precision: Proportion of positive data correctly classified from those predicted as positive by $\hat{f}$:

$prec\left(\hat{f}\right) = \frac{\mathrm{TP}}{\mathrm{TP+FP}}$,

f) Sensibility or recall: It is the hit rate in the positive class also called the true positive rate (TPR): $sens\left(\hat{f}\right) = rec\left(\hat{f}\right) = \mathrm{TPR}\left(\hat{f}\right) = \frac{\mathrm{TP}}{\mathrm{TP+FN}}$,

g) Specificity: It is the hit rate in the negative class. The complement corresponds to the TFP rate: $spec\left(\hat{f}\right) = \frac{\mathrm{TN}}{\mathrm{TN+FP}} = 1 - \mathrm{TFP}\left(\hat{f}\right)$,

The accuracy can be considered as one measure of the model, and the sensitivity or recall as measures of its completeness. The precision and recall are not analyzed separately, but combined into a single measure, which is the weighted harmonic mean of precision and recall (Tafner, 1998):

$$F_1\ \left(\hat{f}\right) = \frac{2 \text{ x prec } \left(\hat{f}\right) \text{ x rec } \left(\hat{f}\right)}{\text{prec } \left(\hat{f}\right) + rec\ \left(\hat{f}\right)}.$$

The ROC curve is a bi-dimensional graph plotted in a called ROC dimension, with (*x,y*) axes representing the rates of false positives (FPR; x) and of true positives (TPR; y). The performance of a

Table 2. Confusion matrix.

Predicted value (obtained by test)		True value (proved by test)	
		Positive	**Negative**
	Positive	TP – True positive	FP – False positive
	Negative	FN – False negative	TN – True negative

Source: Carvalho (2011).

classifier can be plotted to form the curve and represented by a dimensional point (*x, y*) in the ROC dimension (Carvalho, 2011).

In Figure 1, Carvalho (2011) showed a perfect classification by the point (0, 1) where all positive and negative data are correctly classified, and then it is called ROC heaven. Conversely, the point (1, 0) represents the ROC hell. The point (1, 1) are the classifications always positive and the point (0, 0), those always negative. A classifier is considered superior to others when its point is positioned above and to the left of the points of the others classifiers. Given the above, it is common to compare the performance of algorithms in terms of a single measure extracted from the ROC curve: the area under the ROC curve. The area of the ROC curve produces values between 0 and 1; values closer to 1 are considered best, but it is still advisable to calculate the full ROC area through a cross-validation procedure.

The ROC graph used in this study holds the false positives rate (FP Rate) on the horizontal axis (x) and the true positives rate (TP Rate) on the vertical axis (y). The evaluation criterion for a good model of classification is to find the highest true positives rates and the lowest false positive rates in a 0-1 scale. The optimal model is the one that had the ability to hit all data classification which means 100% true positives and 0% false positive (Tahin, 2010).

The ANNs algorithms used in this study were the radial basis function – RBF network, and the multilayers: Multilayer Perceptron and the multilayer perceptronCS (Karaa and Krichene, 2012; Fernandez et al., 2013) and the SVM algorithm selected was the LibSVM (Chang and Lin, 2011). All algorithms in this study belong to the Weka software, which is largely used in data mining and machine learning (Karaa and Krichene, 2012).

The LibSVM is a library of SVM implementations developed by Chang and Lin for several purposes: classification, regression and distribution estimation. The version 3 was used in present study (Chang and Lin, 2011).

The information obtained in the decision tree was also used to build the ANNs. Three ANNs were built, the first with 27 indicators selected from a total of 39 proposed by the PEARLS system; the second was built with 10 indicators selected by market analysts as suitable to analyze the insolvency of credit unions; and the third was made with the R13 indicator from the second neural network added to those 10 indicators, which showed the best performance in building the decision tree.

Eleven PEARLS indicators selected by market analysts as the most adequate to analyze insolvency in credit unions are presented in Table 3. They are: one for protection (P), one for effective financial structure (E), one for asset quality (A), three for rates of return and cost, one for liquidity (L) and four for signals of growth (S).

RESULTS

As previously described, the ANNs were built with 3 algorithms: RBF network, Multilayer perceptron and multilayer perceptronCS. For each algorithm, three ANNs were built: the first with 27 indicators, the second with 10 indicators, and the third with the R13 indicator extracted from the second neural network and added to the group of 10 indicators. For the SVMs model only the LibSVM algorithm was used.

The ANN with RBFNetwork algorithm and 27 indicators

The ANN with 27 indicators and RBF network algorithm obtained a 0.677 Kappa statistic showing that 83.87% of total credit unions were correctly classified by the ANN, while 16.13% were not (Table 4). That means ANN has classified 27 credit unions correctly and 4 incorrectly from the set of 31 solvent credit unions. From the insolvent set, the ANN has classified 25 credit unions correctly and 6 incorrectly (Table 5).

The ANN with multilyer perceptron and multilayer perceptronCS algorithms and 27 indicators

The ANN with multilyer perceptron and multilayer perceptronCS algorithms and 27 indicators had a 0.839 Kappa statistic showing that 91.94% of total credit unions were correctly classified while 8.06% were not (Table 4). That means ANN has classified 27 credit unions correctly and 4 incorrectly from the set of 31 solvent credit unions. From the insolvent set, the ANN classified 30 credit unions correctly and only 1 incorrectly (Table 5).

The SVM with LibSVM algorithm and 27 indicators

The SVM with LibSVM algorithm and 27 indicators obtained a 0.903 Kappa statistic indicating that 95.16% of total credit unions were correctly classified by the SVM, while 4.84% were not (Table 4). The algorithm has classified all the 31 solvent credit unions correctly. From the insolvent set, the SVM classified 28 credit unions correctly and 3 incorrectly (Table 5).

The ANN with RBF network algorithm and 11 indicators

The ANN with RBF network algorithm and 11 indicators had a 0.774 Kappa statistic implying that 88.71% of total credit unions were correctly classified, while 11.29% were

Table 3. Indicators selected by market analysts.

Indicator	Purpose
P1 = Allowance for loan and lease losses / Total Portfolio classified	Measure the amount of allowance for loan and lease losses relative to the total portfolio classified.
E5 = Revenues from financial intermediation / Total Average Assets	Measure the proportion of income from financial intermediation relative to total adjusted assets.
A1 = Fixed assets + Assets not intended to the purpose of cooperative / Adjusted equity	Measure the degree of utilization of own funds with fixed assets and assets not directed to the purpose of the cooperative. The higher the value, the less the focus of the institution on its core business.
R7 = Leftovers / Average Total Assets.	Measure the extent of earnings and also the ability to build social capital. This is an indicator of return on assets.
R8 = Leftovers / Average Adjusted Equity	Measuring the return on equity capital. This is an indicator of return on equity.
R13 = Administrative Costs / Average Total Assets	Measure the percentage of administrative expenses in relation to total assets.
L2 = Short-term Assets / Total Deposits	This indicator is a *proxy* for the liquidity flow.
S1 = Operating Income Growth = (Operating income of the current month / operating income of the previous month) - 1	Measure the rate of growth in operating income.
S6 = Administrative Costs Growth = (administrative cost of the current month / administrative cost of the previous month) -1	Measure the rate of growth of administrative expenses.
S8 = Total Assets Growth = (Total Assets of the current month / total assets of the previous month) -1	Measure the rate of growth of total assets.
S9 = Loan Operations Growth = (Loans of this month / Loans from the previous month) -1	Measure the monthly increase of investments in loans. The higher the index, the more the institution is expanding loan operations.

not (Table 6). The ANN has classified 27 credit unions correctly and 4 incorrectly, from the set of 31 solvent credit unions. From the insolvent set, the ANN classified 28 credit unions correctly and 3 incorrectly (Table 7).

The ANN with multilyer perceptron and multilayer perceptronCS algorithms and 11 indicators

The ANN with multilyer perceptron and multilayer perceptronCS algorithms and 11 indicators had a 0.710 Kappa statistic showing that 85.48% of total credit unions were correctly classified while 14.52% were not (Table 6). The ANN has classified 30 credit unions correctly and only 1 incorrectly, from the set of 31 solvent credit unions. From the insolvent set, ANN also classified 30 credit unions correctly and 1 incorrectly (Table 7).

The SVM with LibSVM algorithm and 11 indicators

The SVM with LibSVM algorithm and 11 indicators had a 0.968 Kappa statistic indicating that 98.39% of total credit unions were correctly classified while 1.61% was not (Table 6). SVM has classified all the 31 solvent credit unions correctly. From the insolvent set, ANN classified 30 credit unions correctly and only 1 incorrectly (Table 7).

The ANN with RBF network algorithm and 10 indicators

The ANN with RBF network algorithm and 10 indicators had a 0.742 Kappa statistic showing that 87.09% of total credit unions were correctly classified, while 12.91% were not (Table 8). That means, ANN has classified 26 credit

Table 4. Summary of the RBF Network, the Multilyer Perceptron and Multilayer PerceptronCS, and the LibSVM algorithms with 27 indicators.

RBF network	Total credit unions	
	No.	%
Correctly classified instances	52	83.87
Incorrectly classified instances	10	16.13
Kappa statistic	0.677	
Multilyer perceptron and multilayer perceptronCS		
Correctly classified instances	57	91.94
Incorrectly classified instances	5	8.06
Kappa statistic	0.839	
LibSVM		
Correctly classified instances	59	95.16
Incorrectly classified instances	3	4.84
Kappa statistic	0.903	

Table 5. The confusion matrix for the RBF network, the multilyer perceptron and the multilayer perceptronCS, and LibSVM algorithms with 27 indicators. Classified as a = INSOLVENT and b = SOLVENT.

Classified as	a	b	<-- classified as
RBF network	27	4	| a = INSOLVENT
	6	25	| b = SOLVENT
Multilyer perceptron and multilayer perceptronCS	27	4	| a = INSOLVENT
	1	30	| b = SOLVENT
LibSVM	31	0	| a = INSOLVENT
	3	28	| b = SOLVENT

unions correctly and 5 incorrectly, from the set of 31 solvent credit unions. From the insolvent set, ANN classified 28 credit unions correctly and 3 incorrectly (Table 9).

The ANN with multilyer perceptron and multilayer perceptronCS algorithms and 10 indicators

The ANN with multilyer perceptron and multilayer perceptronCS algorithms and 10 indicators obtained a 0.710 Kappa statistic indicating that 85.48% of total credit unions were correctly classified while 14.52% were not (Table 8). The ANN has classified 23 credit unions correctly and 8 incorrectly, from the set of 31 solvent credit unions. From the insolvent set, ANN classified 30 credit unions correctly and only 1 incorrectly (Table 9).

The SVM with LibSVM algorithm and 10 indicators

The ANN with LibSVM algorithm and 10 indicators had a 0.968 Kappa statistics indicating that 98.39% of total credit unions were correctly classified, while 1.61% were not (Table 8). SVM has classified all 31 solvent credit unions correctly. From the insolvent set, SVM classified 30 credit unions correctly and only 1 incorrectly (Table 9). According to Table 10, the LibSVM showed superior performance for the error rate of the positive class (amount of false negatives). The same performance is observed when using 10 or 11 indicators, that is, the number of hits within this class does not increase with the presence of the R13 indicator.

The error rate of the negative class which is the portion of the negative class data that were incorrectly classified (amount of false positives) (Table 10). Based on these results, the ANN with the multilayer perceptron and multilayer perceptron algorithms presented superior performance for the three groups of indicators. LibSVM presented the same performance as the multilayer perceptron and multilayer perceptron algorithms for either 10 or 11 indicators (Table 10). Again, the presence of the R13 indicator did not interfere on the evaluation of

Table 6. Summary of the RBF network, the multilyer perceptron and multilayer perceptronCS, and the LibSVM algorithms with 11 indicators.

RBF network	**Total credit unions**	
	No.	**%**
Correctly classified instances	55	88.71
Incorrectly classified instances	7	11.29
Kappa statistic	0.774	
Multilyer perceptron and multilayer perceptronCS		
Correctly classified instances	53	85.48
Incorrectly classified instances	9	14.52
Kappa statistic	0.710	
LibSVM		
Correctly classified instances	61	98.39
Incorrectly classified instances	1	1.61
Kappa statistic	0.968	

Table 7. The Confusion Matrix for the RBFNetwork, the Multilyer Perceptron and the Multilayer PerceptronCS, and LibSVM algorithms with 11 indicators. Classified as a = INSOLVENT and b = SOLVENT.

Classified as	**a**	**b**	**<-- classified as**
RBF network	27	4	| a = INSOLVENT
	3	28	| b = SOLVENT
Multilyer perceptron and multilayer perceptronCS	23	8	| a = INSOLVENT
	1	30	| b = SOLVENT
LibSVM	31	0	| a = INSOLVENT
	1	30	| b = SOLVENT

Table 8. Summary of the RBFNetwork, the Multilyer Perceptron and Multilayer PerceptronCS, and the LibSVM algorithms with 10 indicators.

RBF network	**Total credit unions**	
	No.	**%**
Correctly classified instances	54	87.09%
Incorrectly classified instances	8	12.91%
Kappa statistic	0.742	
Multilyer perceptron and multilayer perceptronCS		
Correctly classified instances	53	85.48%
Incorrectly classified instances	9	14.52%
Kappa statistic	0.710	
LibSVM		
Correctly classified instances	61	98.39%
Incorrectly classified instances	1	1.61%
Kappa statistic	0.968	

insolvency of the credit unions sampled for this study. Table 11 shows the precision rates for the built models. The LibSVM algorithm showed the best results for the three groups of indicators, without difference in precision between the groups with 10 and 11 indicators.

The sensibility of algorithms used for building the models proposed in this study is displayed in Table 11. According to these results, the LibSVM algorithm was superior to others for the three groups of indicators with higher values for the groups of 10 and 11 indicators.

Table 9. The Confusion Matrix for the RBF network, the multilyer perceptron and the multilayer perceptronCS, and LibSVM algorithms with 10 indicators. Classified as a = INSOLVENT and b = SOLVENT.

Classified as	a	b	<-- classified as
RBF network	26	5	| a = INSOLVENT
	3	28	| b = SOLVENT
Multilyer perceptron and multilayer perceptronCS	23	8	| a = INSOLVENT
	1	30	| b = SOLVENT
LibSVM	31	0	| a = INSOLVENT
	1	30	| b = SOLVENT

Table 10. Mean values of TP and FP rates of solvent and insolvent credit unions by the ANN and the LibSVM algorithms.

Indicator	RBF network	Multilayer perceptron and multilayer perceptronCS	LibSVM
	TP rates		
27	0.839	0.919	0.952
11	0.887	0.855	0.984
10	0.871	0.855	0.984
	FP rates		
27	0.194	0.032	0.097
11	0.097	0.032	0.032
10	0.097	0.032	0.032

Table 11. Mean precision, recall, F – Measure (the weighted harmonic mean of precision and recall) and mean ROC area for insolvent and solvent credit unions by the ANN and the LibSVM algorithms.

Indicator	RBF network	Multilayer perceptron and multilayer perceptronCS	LibSVM
	Mean precision		
27	0.841	0.923	0.956
11	0.888	0.874	0.984
10	0.873	0.874	0.984
	Mean recall		
27	0.839	0.919	0.952
11	0.887	0.855	0.984
10	0.871	0.855	0.984
	F – Measure		
27	0.839	0.919	0.951
11	0.887	0.853	0.984
10	0.871	0.853	0.984
	Mean ROC area		
27	0.869	0.947	0.952
11	0.879	0.870	0.984
10	0.889	0.879	0.984

As previously described in methods, precision and sensibility should not be analyzed separately, but jointly by the F- Measure (the weighted harmonic mean of precision and recall). As shown in Table 11, the LibSVM algorithm presented the highest values for the three groups of indicators with superior result when working with 10 or 11 indicators instead of 27 indicators.

The ROC area as evidenced by the ROC graph analysis is larger by the LibSVM algorithm for the three groups of indicators (Table 11). The groups containing either 10 or 11 indicators showed the highest LibSVM value.

DISCUSSION

SVM model was included in this work because currently is one of the machine-learning algorithms most used to evaluate binary systems suitably. The algorithms of the ANNs and SVMs were used to build models of classification suited to select one financial standard for identifying status of insolvency of credit unions, bank and life insures (Hsiao and Whang, 2009; Karra and Krichene, 2012; Ribeiro et al., 2012). To this end, we used some indicators proposed by the PEARLS system, which were grouped into 3 sets of 27, 11 and 10 indicators previously defined.

Comparing the results of ANNs and SVMs, it is evident that superiority of the SVMs as binary classifier of solvency since its LibSVM algorithm showed the best results in all evaluations of performance proposed in this study, except for the error rate of the negative class which indicates the incorrect classifications in the negative class.

From the ROC curve, it was possible to observe the highest TP and FP rates for the ANNs algorithms, which resulted in more liberal models. On the other hand, the LibSVM algorithm generated more conservative models since it showed good performance with respect to the FP rates, but few high TP rates. The performance of the ANNs by the multilayer perceptron, multilayer perceptronCS and the RBF network algorithms in classifying data was inferior to the LibSVM.

Even though only with a single ANN algorithm, the performance would be better likely by classifying a new credit union as true positive (INSOLVENT). It can be deduced just by observing separately the performance curves in the ROC graph.

It is still early to say that the ANNs and the SVMs algorithms can replace other methods for evaluating insolvency. However, it can be still helpful as tools of support for detecting signs and risks of imminent insolvency in credit unions and banks (Karaa and Krichene, 2012).

With respect to the number of indicators of the PEARLS system, appropriate to evaluate insolvency of credit unions, there is no need to use all those 39 indicators previously proposed. It can be confirmed by the same high efficiency of the group with 10 indicators elected by market analysts for this evaluation. And still, adding the R13 indicator, prominently in the decision tree, to the group of the 10 indicators did not alter the amount of hits by the models proposed in this study.

Conclusions

The major implication of this work is reducing the gap existing on this theme by suggesting new financial tools to evaluate the health of credit unions which are very important to the socioeconomic development of the regions where they are located.

This work proposes to help credit unions to meet solid financial structures in order to comply with its business mission, valuing relationships, offering financial solutions for adding income and improving the life quality of their members and the society.

Conflict of Interests

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENT

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

REFERENCES

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

Azmathullah HM, Deo MC, Deolalikar PB (2005). Neural networks for estimation of scour downstream of a ski-jump bucket. J. Hydraul. Eng. 131:898–908. http://dx.doi.org/10.1061/(ASCE)0733-9429(2005)131:10(898)

Azmathullah HM, Deo MC, Deolalikar PB (2006). Estimation of scour below spillways using neural networks. J. Hydraul. Res. 44:61–69. http://dx.doi.org/10.1080/00221686.2006.9521661

Azmathullah HM, Deo MC, Deolalikar PB (2008) Alternative neural networks to estimate the scour below spillways. Adv. Eng. Softw. 39:689–698. http://dx.doi.org/10.1016/j.advengsoft.2007.07.004

Azmathullah HM, Wu FC (2011). Support vector machine approach to for longitudinal dispersion coefficients in streams. Appl. Soft. Comput. 11:2902–2905. http://dx.doi.org/10.1016/j.asoc.2010.11.026

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. P. 123.

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

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. http://dx.doi.org/10.5700/rausp1011

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
Carvalho A (2011). Inteligência Artificial – uma abordagem de aprendizado de máquina. 1º ed. LTC, São Paulo, Brazil. P. 394.
Chang CC, Lin CJ (2011). LIBSVM: A library for support vector machines. T. Intel. Syst. Technol. 2:1–27. http://140.112.30.28/~cjlin/papers/libsvm.pdf
Fernandez FG, Santos ISL, Martinez JL, Izquierdo SI, Redondo FL (2013). Use of artificial neural networks to predict the business success or failure of start-up firms. Inetch, pp. 245–256. http://dx.doi.org/10.5772/51381
Ferreira FS (2010). Aplicação da técnica de regressão logística na previsão de insolvência no segmento das cooperativas de crédito no Brasil. Universidade Federal do Rio de Janeiro, Brazil. MSc Dissertation. P. 127.
Han J, Kamber M (2006). Data mining: concepts and techniques. 2º ed. University of Illinois at Urbana-Champaign. Elsevier, New York, P. 703.
Hsiao SH, Whang TJ (2009). A study of financial insolvency prediction model for life insurers. Expert Syst. Appl. 36:6100–6107. http://dx.doi.org/10.1016/j.eswa.2008.07.024
Jardin P (2010). Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy. Neurocomputing 73:2047–2060. http://dx.doi.org/10.1016/j.neucom.2009.11.034
Karra A, Krichene A (2012). Credit-risk assessment using support vector machine and multilayer neural network models: a comparative study of Tunisian bank. Acc. Manage. Infor. Syst. 11:587–620. http://www.cig.ase.ro/RePEc/ami/articles/11_4_5.pdf
Kohonen T (1988). An introduction to neural computing. Neural Networks. 1:3–16. http://dx.doi.org/10.1016/0893-6080(88)90020-2
Monard MC, Baranauskas JA (2003). Conceitos sobre aprendizado de máquina. In: Rezende SO (Ed.). Sistemas Inteligentes: fundamentos e aplicações. São Carlos: Manole. pp. 89–114.
Pinheiro MAH (2003). Gerenciamento do risco de crédito: um modelo para previsão de insolvência em cooperativas de crédito. Instituto de Educação Continuada da PUC-MG, Minas Gerais, Brazil. Monography.
Ribeiro B, Silva C, Chen N, Vieira A, Neves JC (2012). Enhanced default risk models with SVM+. Expert Syst. Appl. 39:10140–10152. http://dx.doi.org/10.1016/j.eswa.2012.02.142
Richardson DC (2002). PEARLS Monitoring System. Word Council of Credit Unions. Toolkit series number 4. October, 2002. Retrieved from <http://www.woccu.org/best practices. Accessed on 20 Oct. 2011.
Tafner MA (1998). Redes Neurais Artificiais: aprendizado e plasticidade. Cérebro & Mente, 2(5). Online. http://www.cerebromente.org.br/n05/tecnologia/plasticidade2.html
Tahin APN (2010). Máquinas de vetores de suporte (SVM). Tutorial técnico, UFSC, Florianópolis. Retrieved from: <http://www.andretahoim.com>. Accessed on 08 Jul. 2012.
WORD COUNCIL OF CREDITS UNIONS – WOCCU. PEARLS Monitoring System. (2002). Retrieved from: <http://www.woccu.org/publications/researchpub>. Accessed on 05 Jan. 2011.
Wu CC, Tzeng GH, Goo YJ, Fang WC (2007). A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy. Expert Syst. Appl. 32:397–408. http://dx.doi.org/10.1016/j.eswa.2005.12.008