Expert Systems with Applications 40 (2013) 356–364

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Credit scoring models for the microfinance industry using neural networks: Evidence from Peru

Antonio Blanco [a,*], Rafael Pino-Mejías [b], Juan Lara [c], Salvador Rayo [c]

[a] Department of Financial Economics and Operations Management, Faculty of Economics and Business Studies, University of Seville, Avda. Ramon y Cajal, 1, 41018 Seville, Spain
[b] Department of Statistics and Operational Research, Faculty of Mathematics, University of Seville, Avda. Reina Mercedes, s/n 41012 Seville, Spain
[c] Department of Financial Economics and Accounting, Faculty of Economics and Business Studies, University of Granada, Campus Cartuja, s/n 18071 Granada, Spain

ARTICLE INFO

Keywords:
Microfinance institutions
Classification rules
Multilayer perceptron
Linear discriminant analysis
Quadratic discriminant analysis
Logistic regression

ABSTRACT

Credit scoring systems are currently in common use by numerous financial institutions worldwide. However, credit scoring with the microfinance industry is a relatively recent application, and no model which employs a non-parametric statistical technique has yet, to the best of our knowledge, been published. This lack is surprising since the implementation of credit scoring should contribute towards the efficiency of microfinance institutions, thereby improving their competitiveness in an increasingly constrained environment. This paper builds several non-parametric credit scoring models based on the multilayer perceptron approach (MLP) and benchmarks their performance against other models which employ the traditional linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and logistic regression (LR) techniques. Based on a sample of almost 5500 borrowers from a Peruvian microfinance institution, the results reveal that neural network models outperform the other three classic techniques both in terms of area under the receiver-operating characteristic curve (AUC) and as misclassification costs.

1. Introduction

Over the last decade, the microfinance sector[1] has grown dramatically, and is currently considered as a booming industry. In the period 1998–2008, the number of microfinancial institutions (hereinafter, MFIs) grew by 474%, and the number of customers increased by 1048%. Attracted by this rapid growth, a large number of international commercial banks have started operating in the microfinance sector, viewing it as a potential for profitable investment. This injection of interest has increased the competition between the players in this industry, and has negatively affected the MFIs. The MFIs therefore need to increase their efficiency in all their processes, minimize their costs, and control their credit risk if they want to survive in the long-term. One way for the MFIs to become more efficient in order to compete with the commercial banks is the implementation of automatic credit scoring[2] systems to evaluate their credit applicants since credit scoring reduces the cost of credit analysis, improves cash flow, enables faster credit decisions, reduces the losses, and also results in the closer monitoring of existing accounts and the prioritization of repayment collection. To this end, Rhyne and Christen (1999) suggest that credit scoring is one of the most important uses of technology that may affect microfinance, and Schreiner (2004) affirms that experiments carried out in Bolivia and Colombia show that the implementation of credit scoring improves the judgment of credit risk and thus cuts, by more than $75,000 per year, the costs of MFIs. Nevertheless, and in contrast to the concentration of research on financial institutions, the development of credit scoring models in the microfinance sector has only undergone minor advances. Furthermore, those models in existence are based on traditional parametric statistical techniques, mainly linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and logistic regression (LR), despite the overwhelming evidence found in numerous studies which indicates that the non-parametric methodologies usually outperform these classic statistical models (for example, see Lee & Chen, 2005; West, 2000). That is, to the best of the authors' knowledge, in the existing literature

* Corresponding author. Tel.: +34 954 559 875; fax: +34 954 557 570.
E-mail address: aj_blanco@us.es (A. Blanco).

[1] In the microfinance sector, operate the Microfinance institutions (hereinafter MFIs) which offer savings services and small loans (namely microcredits) to those sectors of the population with the greatest problems of access to financial resources. Therefore, the MFIs exercise relevant social work since they financially support the poorest people, who, by creating a microenterprise, can escape the socioeconomic situation of exclusion in which they find themselves. For this reason, the goals and management criteria of the many MFIs lay less emphasis on business components and greater emphasis on social components than those used by their new competitors (international commercial banks).

[2] The objective of credit scoring models is to assign credit applicants to one of two groups: either to a 'good credit' group that is likely to repay the financial obligation or a 'bad credit' group that should be denied credit because of a high likelihood of defaulting on the financial obligation (Hand & Henley, 1997).

0957-4174/$ - see front matter
http://dx.doi.org/10.1016/j.eswa.2012.07.051

no credit scoring model designed for the microfinance industry applies a non-parametric methodology, and therefore, the microfinance industry has not yet benefited from the advantages of non-parametric techniques to improve the performance of credit scoring models, and hence are failing to compete on equal terms with their new competitors, the international commercial banks. Of the few credit scoring models developed for MFIs, all have used parametric methodologies, particularly LDA and LR (Kleimeier & Dinh, 2007; Rayo, Lara, & Camino, 2010; Reinke, 1998; Sharma & Zeller, 1997; Viganò, 1993; Vogelgesang, 2003; Zeller, 1998). However, the strict assumptions (linearity, normality and independence among predictor variables) of these traditional statistical models, together with the pre-existing functional form relating response variables to predictor variables, limit their application in the real world. Several authors (for example, Karels & Prakash, 1987; Reichert, Cho, & Wagner, 1983) point out that two basic assumptions of LDA are often violated when applied to credit scoring problems: (a) the independent variables included in the model are multivariate and normally distributed, (b) the group dispersion matrices (or variance–covariance matrices) are equal across the failing and the non-failing groups (for a detailed analysis of the problems in applying discriminant analysis in credit scoring models, see Eisenbeis, 1978). In the cases where the covariance matrices of the two populations are unequal, theoretically, QDA should be adopted, although LDA is reported to be a more robust and precise technique (Dillon & Goldstein, 1984). In the same way as LDA, LR is also optimal under the assumption of multivariate normal distributions with equal covariance matrices, and LR also remains optimal in a wider variety of situations. However, logistic regression requires larger data sets to obtain stable results, interactions between predictor variables must be formulated, and complex non-linear relations between the dependent and independent variables could be incorporated through appropriate but not evident transformations. For these reasons, in recent years, non-parametric statistical models, such as the *k*-nearest neighbor algorithm (Henley & Hand, 1996), support vector machines (Vapnik, 1998), decision tree models (Davis, Edelman, & Gammerman, 1992), and neural network models (Patuwo, Michael, & Ming, 1993), have been successfully applied to credit scoring problems. Of these, artificial neural networks (ANNs) constitute one of the most powerful tools for pattern classification due to their non-linear and non-parametric adaptive-learning properties. Many studies have been conducted that have compared ANNs with other traditional classification techniques in the field of credit scoring models, since the default prediction accuracies of ANNs are better than those using classic LDA and LR (Arminger, Enache, & Bonne, 1997; Desai, Conway, Crook, & Overstreet, 1997; Desai, Crook, & Overstreet, 1996; Hand & Henley, 1997; Lee & Chen, 2005; Lee, Chiu, Lu, & Chen, 2002; Malhotra & Malhotra, 2002; Markham & Ragsdale, 1995; Patuwo et al., 1993; Piramuthu, 1999; Srinivasan & Ruparel, 1990; West, 2000). However, despite yielding satisfactory results, ANNs also feature certain disadvantages, such as its black box nature and the long training process involved in the design of the optimal network topology (Chung & Gray, 1999).

The main goal of this paper is therefore to develop a credit scoring model specially designed for the microfinance industry by using multilayer perceptron neural networks (hereinafter, MLP). Moreover, we also compare the performance of MLP models against the three parametric techniques most widely used: linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and logistic regression (LR). Based on a large sample which contains financial and non-financial variables of almost 5500 borrowers from a Peruvian MFI, seventeen credit scoring models are created, of which fourteen are MLP-based models.

The remainder of our paper proceeds as follows. In Section 2, details of our data set are provided, and a detailed examination of the variables available is undertaken in order to predict the default. In Section 3, several credit scoring models specifically designed for MFIs are developed. To his end, various methodologies are employed: Fisher discriminant analysis, logistic regression, and multilayer perceptron. In Section 4, the results of different models are shown and their comparison is made. An extensive discussion on the results is also carried out. Finally, Section 5 provides the main conclusion of this study and future research lines are analyzed.

2. Data and variables

2.1. The data set

We use a data set of microcredits from a Peruvian Microfinance Institution (*Edpyme Proempresa*). Our dataset contains customer information during the period 2003–2008 related to: (a) personal characteristics (marital status, gender, etc.); (b) economic and financial ratios of their microenterprise; (c) characteristics of the current financial operation (type interest, amount, etc.); (d) variables related to the macroeconomic context; and (e) any delays in the payment of a microcredit fee. After eliminating missing and abnormal cases, 5451 cases remain. From among these, 2673 (49.03%) are default cases, and 2778 (50.97%) are not. In line with other studies (for example, Schreiner, 2004), a microcredit presenting a delay in repayment of at least fifteen days is defined as default microcredit. To perform an appropriate comparison of the classification models, (LDA, QDA, LP, and MLP), our final data set is randomly split into two disjoint sub-sets; a training set of 75% and a test set of 25%. The test sample contains a total of 1363 cases (51.80% failed and 48.20% non-failed). The configuration of parameters of each model is selected through a 10-fold cross-validation procedure, as described in Sections 3.1–3.3. One advantage of cross-validation is that the credit scoring model is developed with a large proportion of the available data (75% in this case).

2.2. Description of input variables

Table 1 shows the input variables used in this study.[3] They provide the various characteristics of borrowers, lenders, and loans. Numerous qualitative variables are considered in our study, since: (a) Schreiner (2004) suggests that the input variables of the credit scoring forces the microfinance sector to be more qualitative and informal than those considered by traditional banks; and (b) recent literature concludes that the inclusion of qualitative variables improves the prediction power of models. Moreover, since the default of borrowers has a close relationship with the general economic situation, variables linked to the macroeconomic context are also considered as input variables. With respect to the dependent variable, default of the microcredit, this takes a value of 1 if the microcredit fails, and 0 otherwise.

The first ratio indicates the number of times the income exceeds total assets. Therefore, we estimate that the ratio ($R1$) is inversely related with respect to the probability of default. The ratio $R2$ measures the relationship between the gross and operating costs of the microenterprise. As with the previous ratio, we expect that the sign of its coefficient is negative since the higher the value of this ratio, the more solvent the income/loss of the firm, and the lower the financial difficulties. The third financial ratio ($R3$) measures the liquidity of the microenterprise. Due to the design of this ratio, the higher its value, the lower the probability of default. Therefore, the sign of the estimator is expected to be negative. The fourth

[3] This table also shows the expected sign of the relationship between each input variable and the probability of default. The statistical descriptions of all the input variables are shown in Table 1 and Table 2 of Appendix 1. These statistics are presented for each group (failed and non-failed).