ELSEVIER

Expert Systems with Applications 28 (2005) 743–752

Expert Systems with Applications

www.elsevier.com/locate/eswa

A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines

Tian-Shyug Lee[a,*], I-Fei Chen[b]

[a]Graduate Institute of Management, Fu-Jen Catholic University, No. 510, Chung-Cheng Road, Hsinchuang, Hsin-Chuang, Taipei 24205, Taiwan, ROC
[b]Graduate Institute of Business Administration, Fu-Jen Catholic University, Hsin-Chuang, Taipei, Taiwan, ROC

Abstract

The objective of the proposed study is to explore the performance of credit scoring using a two-stage hybrid modeling procedure with artificial neural networks and multivariate adaptive regression splines (MARS). The rationale under the analyses is firstly to use MARS in building the credit scoring model, the obtained significant variables are then served as the input nodes of the neural networks model. To demonstrate the effectiveness and feasibility of the proposed modeling procedure, credit scoring tasks are performed on one bank housing loan dataset using cross-validation approach. As the results reveal, the proposed hybrid approach outperforms the results using discriminant analysis, logistic regression, artificial neural networks and MARS and hence provides an alternative in handling credit scoring tasks.

Keywords: Credit scoring; Classification; Neural networks; Multivariate adaptive regression splines; Cross-validation

1. Introduction

Credit risk evaluation decisions are crucial for financial institutions due to high risks associated with inappropriate credit decisions that may result in huge amount of losses. It is an even more important task today as financial institutions have been experiencing serious challenges and competition during the past decade. When considering the case regarding the application for a large loan, such as a mortgage or a construction loan, the lender tends to use the direct and individual scrutiny by a loan officer or even a committee. However, if hundreds of thousands, even millions of credit card or consumer loan applications need to be evaluated, the financial institutions will usually adopt models to assign scores to applicants rather than examining each one in detail. Hence various credit scoring models need to be developed for the purpose of efficient credit approval decisions.

With the tremendous growth of the credit industry and the diversified loan portfolios nowadays, credit scoring has gained more and more attention as the credit industry can then benefit from on time decisions, reducing possible risks, improving cash flow, and insuring proper credit collections. Aiming to satisfy the above-mentioned needs, many different useful techniques, known as the credit scoring models, have been developed by financial institutions and researchers in order to solve the problems involved during the evaluation process. The objective of credit scoring models is to assign credit applicants to either a 'good credit' group that is likely to repay financial obligation or a 'bad credit' group, with high possibility of defaulting on the financial obligation, whose application should be denied. Therefore credit scoring lies in the domain of the more general and widely discussed classification problems (Anderson, 1984; Dillon & Goldstein, 1984; Johnson & Wichern, 2002). The classification problems where items/observations can be assigned to one of several known disjoint groups have long played important roles in business related decision making due to its wide applications in decision support, financial forecasting, fraud detection, marketing strategy, process control, and other related fields (Cabena, Hadjinaian, Stadler, Verhees, & Zanasi, 1997, Chen et al., 1996, Fayyad, Piatetsky-Shapiro, & Smyth, 1996).

Usually, credit scoring is applied to rank credit information based on the application form details and

* Corresponding author. Tel.: +886 2 2903 1111x2905; fax: +886 2 2901 8475.
E-mail address: badm1004@mails.fju.edu.tw (T.-S. Lee).

0957-4174/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2004.12.031

other relevant information held by a credit reference agency. As the results, accounts with high possibility of default can be monitored and necessary actions can be taken in order to prevent the account from being default. In response, the statistical methods, non-parametric methods, and artificial intelligence approaches have been proposed to support the credit approval decision process (Desai, Crook, & Overstreet, 1996; Lee, Chiu, Lu, & Chen, 2002; Thomas, 2000; West, 2000).

After careful review of the crediting scoring literature, it can be concluded that linear discriminant analysis (LDA) and logistic regression were the two most commonly used statistical techniques in building credit scoring models. However, the utilization of linear discriminant analysis has often been criticized due to the assumptions of linear relationship between dependent and independent variables, which seldom holds, and the fact that it is sensitive to deviations from the multivariate normality assumption (Karels & Prakash, 1987; Reichert, Cho, & Wagner, 1983). Theoretically, quadratic discriminant analysis (QDA) should be adopted when the covariance matrices of the different populations are unequal. However, QDA seems to be more sensitive to the model assumptions than LDA and LDA has reported to be a more robust and precise method (Dillon & Goldstein, 1984; Sharma, 1996).[1] In addition to the LDA approach, logistic regression is another commonly utilized alternative to conduct credit scoring tasks. Basically, the logistic regression model was emerged as the technique in predicting dichotomous outcomes. Logistic regression does not require the multivariate normality assumption, however, the dependent variable exposed to a full linear relationship among independent variables in the exponent of the logistic function. Basically, both LDA and logistic regression are designed for the case when the underlying relationship between variables are linear and hence are reported to be lack of enough credit scoring accuracy (Thomas, 2000; West, 2000).

Artificial neural networks provide a new alternative to LDA and logistic regression in handling credit scoring tasks, particularly in situations where the dependent and independent variables exhibit complex non-linear relationships. Even though neural networks have reported to provide better credit scoring accuracy than those using LDA and logistic regression (Desai et al., 1996; Jensen, 1992; Lee et al., 2002; Piramuthu, 1999; West, 2000), it is, however, also being criticized for its long training process in obtaining the optimal network's topology, not easy to identify the relative importance of potential input variables, and certain interpretive difficulties and hence has limited its applicability in handling general classification and credit scoring problems (Craven & Shavlik, 1997; Lee et al., 2002; Piramuthu, 1999).

In addition to the above-mentioned techniques, multivariate adaptive regression splines (MARS) is another commonly discussed classification technique nowadays (Friedman, 1991). MARS is widely accepted by researchers and practitioners for the following reasons. Firstly, without the drawbacks of LDA and logistic regression, MARS is capable of modeling complex non-linear relationship among variables without strong model assumptions. On the other hand, unlike neural networks, MARS can capture the relative importance of independent variables to the dependent variable when many potential independent variables are considered. Thirdly, MARS does not need long training process and hence can save lots of model building time, especially when the dataset is huge. Finally, one strong advantage of MARS over other classification techniques is the resulting model can be easily interpreted. It not only points out which variables are important in classifying objects/observations, but also indicates a particular object/observation belongs to a specific class when the built rules are satisfied. The final fact has important managerial and interpretative implications and can help to make appropriate decisions.

Based on the above-mentioned modeling advantages of MARS, the authors believe that MARS should be a good supporting tool for neural networks as the technical merits of MARS are just the shortcomings of neural networks. Using MARS as a first-stage modeling tool with the obtained results being the inputs to neural networks should contribute to the success of the subsequent model building tasks. Focusing on improving the above-mentioned drawbacks of neural networks credit scoring models, the purpose of this study is to explore the performance of credit scoring with a two-stage hybrid modeling procedure using artificial neural networks and multivariate adaptive regression splines (MARS). The rationale underlying the analyses is firstly to use MARS in building the scoring model, the obtained significant variables are then used as the input variables of the designed neural networks model. Please note that, according to the knowledge of the authors, there still does not exist a theoretical method, which can optimally determine the appropriate input nodes of a neural networks model; MARS can be implemented as a generally accepted method for identifying important variables when many potential independent variables are considered. Finally, as the two-stage modeling procedure will use the obtained significant variables from MARS as input nodes, hence it can reduce the number of input nodes, simply the network structure, and shorten the model building time.

To demonstrate the feasibility and effectiveness of the proposed two-stage credit scoring procedure, credit scoring tasks are performed on one housing loan dataset. As cross-validation is the preferred procedure in testing the out-of-sample classification capability of the built classification model (Breiman, Friedman, Olshen, & Stone, 1984;

[1] Since LDA has reported to be a more robust method than QDA when the theoretical presumptions are violated, hence the LDA approach will be used in building the credit scoring model in this study.