Expert Systems with Applications 36 (2009) 8333–8341

Contents lists available at ScienceDirect

Expert Systems with Applications

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

A new two-stage hybrid approach of credit risk in banking industry

Shu Ling Lin *

Department of Business Management, College of Management, National Taipei University of Technology, No. 1, Three section, Chung Hsiao East Road, Taipei 10608, Taiwan, ROC

ARTICLE INFO

JEL Classification:
C45
G21
G32

Keywords:
Credit risk modeling
Financial distress
Logistic regression
Artificial neutral network (ANN)
Two-stage hybrid model

ABSTRACT

The rationale under the analyses is to propose a new approach by three kinds of two-stage hybrid models of logistic regression-ANN, to explore if the two-stage hybrid model outperformed the traditional ones, and to construct a financial distress warning system for banking industry in Taiwan. The differences from the literatures are that this study adopts the "optimal cutoff point" approach proposed by Hosmer and Lemeshow [Hosmer, D. W., & Lemeshow, S. L. (2000). *Applied logistic regression* (2nd ed.). New York: A Wiley-Interscience], to determine the cutoff point for financial distress. Additionally, cross-validation [Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman and Hall; Stone, M. (1974). Cross-validation choice and assessment of statistical predictions. *Journal of Royal Statistical Society. Series B*, 36, 111–147] is used to evaluate the prediction power of the constructed models.

The results find the factors of *observable loans to total loans*, *allowance for doubtful accounts recovery rate*, and *interest-sensitive assets to liabilities ratio* are significantly related to the financial distress of banks in Taiwan. In the prediction of financially distressed, two-stage hybrid model giving the best performance of 80.0% using cross-validation approach and demonstrates stronger prediction power than conventional logistic regression, logarithm logistic regression, and ANN approaches. It demonstrates that the two-stage hybrid model outperforms the conventional method, providing an alternative in handling credit risk modeling which have assessment implications for analysts, practitioners, and regulators.

1. Introduction

The assessment of the financial distress warning system of banks is an important issue for supervising authorities and investors. Financial institutions play an important financial intermediation role in the economy; the development of the financial sector influences the level of the economy's development. Credit risk[1] induces financial distress on banks, and its assessment requires advanced modeling techniques that will link it to the sources of uncertainty generated. Accordingly, credit risk remains one of the major threats that financial institutions face, and it is essential to model the credit risk of financial institutions. The Basel committee on Banking Supervision proposes a capital adequacy frame that allows banks to measure capital requirement for their banking books using internal assessments of key risk drivers. Thus, there is necessity for systems to assess credit risk for banks.

In the moves to liberalize and internationalize the financial markets, Taiwan deregulated the banking industry in 1991 by allowing the establishment of new-privatized banks and encouraged the establishment of financial holding companies in 2002. In the face of increasingly intense competition, banks opt to shift business focus from business finance to consumer finance. However, aggressive and practically unchecked marketing of consumer lending has led to swelling of non-performing loan problems and has caused financial distress in some banks, which in turn creates economic and social turmoil. In the midst of a rapidly changing financial environment, a few poorly run banks fell into financial distress which created considerable adverse impact on the economy as a whole. The United States secondary mortgage turmoil caused by the global credit crisis is a best example. Therefore, banks face the necessity of managing financial distress, and its management has distinguished qualities due to higher exposure to credit risk by the role of intermediaries they play on the financial markets. The above phenomenon implies an appropriate financial distress warning system established which is most urgent of issues.

Previous researches on credit risk modeling refers to traditional multivariate statistical and econometric techniques, for instance multivariate discriminant analysis, logit and probit models are among the most widespread for credit risk modeling (Altman, Avery, Eisenbeis, & Stinkey, 1981; Altman & Saunders, 1998; Beaver, 1966; Gordy, 2000; Lin & Chang, 2006), as well

* Tel.: +886 2 29052664; fax: +886 2 29019779.
E-mail addresses: trad1008@mails.fju.edu.tw, shundy@ms24.hinet.net.

[1] According to the Basel Committee on Banking Supervision, Bank of International Settlements (2000) defined the credit risk as the potential risk that a bank borrower or counterparty will fail to meet it obligations in accordance with agreed terms.

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.10.015

methodologies from the field of artificial intelligence and operational research (Dimitras, Zanakis, & Zopounidis 1998; Zopounidis & Doumpos, 2002). Over the past three decades, more sophisticated methodologies have been developed to support corporate financial distress assessment which included gambler ruin approach and option pricing model. The theoretical foundation about how a firm gets bankrupted when the liquidation value of its assets falls below its debt obligations. Such model was developed and modified by Merton (1974); Scott (1981) and Wilcox (1972). Recent studies have considered new non-parametric methods such as mathematical programming, classification trees, neural networks (Abdou, Pointon, & Elmasry, 2008; Angelini, Tollo, & Roli A., 2008; Yeh & Lien, 2008; Yu, Wang, & Lai, 2008) and support vector machines. As Baesens et al. (2003) concluded that credit risk data are weakly non-linear, thus suggesting that highly complex or non-linear models are not expected to have a considerable predicting ability compared to simpler models.

Most of the existing studies on the development of credit risk models involve non-financial firms. However, financial institutions, such as banks, investment firms, security brokers, etc., are also apt to default and major interest due to their crucial role in the economic and business activity in financial markets. Therefore, developing credit-rating models for financial institutions, especially for banks, is of major interest to analysts, practitioners, and supervising authorities. Sahajwala and Van den Bergh (2000) review the current practices on the development and use of credit-rating models for banks.

On the quantitative approach of financial distress warning system, Demirgüç-Kunt & Detragiache, 1998 use multivariate logit approach to establish an early warning model for banking crisis by sampling the data of developed and developing countries from 1980 to 1994. Their study finds that banking crisis tends to occur when the macroeconomic environment is weak; especially when the growth of GDP is low, inflation is high, and when it is coupled with increase in banking sector risk. Johnson, Boone, Breach, and Friedman (2000) posit that in the 1997 Asian financial crisis, the measure of corporate governance has better explanatory power than macroeconomic measures and in countries with weak corporate governance, financial distress problems are more likely to emerge.

Claessens, Djankov, and Lang (1999); Claessens, Djankov, Fan and Lang (1999) and La Porta, Lopez-de-Silanes, and Shleifer (1999) also find that many listed firms use the schemes of pyramid structures and cross-holding to gain corporate control and the financial distress of a firm is to a certain extent related to the cross-holding between parent company and subsidiaries and poor management, and not entirely attributed to financial factors. Thus, to effectively reduce the likelihood of financial distress, the supervision mechanism of board of directors needs to come into play. The nonmonotonic view of Morck, Shleifer, and Vishny (1988) proposes that based on entrenchment hypothesis and convergence-of-interests hypothesis, low ownership by the board of directors limits the board's ability and willingness to supervise the management and might be positively related to the likelihood of financial distress. As shown by the discussions above, corporate governance plays a vital role in the financial soundness of a firm. Therefore, this study also includes the variable of ownership by directors and supervisors to discern whether higher shareholding by directors and supervisors would increase or decrease the likelihood of financial distress. Lee, Chiu, and Lu (2002) and Lee and Chen (2005) explore the performance of credit scoring using a two-stage hybrid modeling procedure with artificial neural networks and multivariate adaptive regression splines (MARS). 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.

Consequently, this study proposes three two-stage hybrid models of logistic regression-ANN, compare with logistic regression, Altman's (2006) logarithm logistic regression, artificial neutral network (ANN), aims to explore if the two-stage hybrid model outperformed the conventional ones, to construct a financial distress warning system suitable for Taiwan's banking industry, and to provide an optimal model of credit risk for supervising authorities, analysts and practitioners in conducting risk assessment and decision making. The differences from the literatures are that this study adopts the "optimal cutoff point" approach proposed by Hosmer and Lemeshow (2000), to determine the cutoff point for financial distress. In addition, cross-validation (Efron & Tibshirani, 1993; Stone, 1974) is used to evaluate the prediction power of the constructed models. The objective of this study is to construct a financial distress warning model of banks in Taiwan, in distinguishing financially sound and financially distressed banks by the criteria of capital adequacy ratio below 8%, loss of capital by one third or more, current ratio below 7%, non-performing loan ratio over 5%, or delisted declared during the 2002–2004 period. In order to improve investor awareness of credit risk, prior to establishing the financial distress early warning system, detection advance firms may change the deal to reduce the preventable loss for stakeholders.

The remainder of this paper is structured as follows. Section 2 discusses the model and research design issues. Section 3 presents and analyzes the data description and results while the final section concludes with a summary of results.

2. Methodology

2.1. Two-stage model of logistic regression and ANN

Prior studies show that logistic regression has better prediction power. Altman (2006) proposes a logarithm logistic regression approach that takes the log of independent variables to enhance prediction accuracy. Artificial neural network (ANN) is a unique statistical technique that has massive computing power, powerful memory, learning ability and fault tolerance ability. According to Lee and Chen (2005), the use of two-stage hybrid model enhances accuracy. Thus this study employs logistic regression, logarithm logistic regression, ANN, and two-stage models for the development of financial distress models, analysis and comparison.

This study integrates the framework of logistic regression and ANN and proposed three kinds of two-stage hybrid model. In stage I, influencing variables are selected using logistic regression. In stage II, the influencing variables are taken as the input variables of BPN. It is hoped that by providing the ANN with a good starting point, a more precise model can be developed on the strength of its learning ability. Such a hybrid model is then compared with logistic regression, logarithm logistic regression and ANN.

2.1.1. Two-stage model of logistic regression and ANN (1)

Firstly, substitute dependent variable Y and independent variables $X_1, X_2, \ldots, X_p$ into logistic regression. Use logistic regression with Wald-forward method to identify independent variables with significant influence on distress probability $X_1^*, X_2^*, \ldots, X_k^*$.

Secondary, substitute significant variables $X_1^*, X_2^*, \ldots, X_k^*$ in ANN model as independent variables of the input layer, and Y as dependent variable of the input layer to obtain a set of predictive values for distress probability ($\hat{P}$). Find the financial distress cutoff point for predictive values for distress probability ($\hat{P}$) based on the point of intersection of sensitivity and specificity according to Hosmer and Lemeshow (2000) and compare the results with actual values (Y). The ANN model and the cutoffs can be used for prediction based on other datasets.