Expert Systems with Applications 36 (2009) 4075–4086

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Using neural networks and data mining techniques for the financial distress prediction model

Wei-Sen Chen *, Yin-Kuan Du

Industrial Technology Research Institute, #195, Sec. 4, Chung-Hsing Rd., Chutung 310, HsinChu, Taiwan, ROC

ARTICLE INFO

Keywords:
Financial distress prediction model
Artificial neural network
Data mining

ABSTRACT

The operating status of an enterprise is disclosed periodically in a financial statement. As a result, investors usually only get information about the financial distress a company may be in after the formal financial statement has been published. If company executives intentionally package financial statements with the purpose of hiding the actual status of the company, then investors will have even less chance of obtaining the real financial information. For example, a company can manipulate its current ratio by up to 200% so that its liquidity deficiency will not show up as a financial distress in the short run. To improve the accuracy of the financial distress prediction model, this paper adopted the operating rules of the Taiwan stock exchange corporation (TSEC) which were violated by those companies that were subsequently stopped and suspended, as the range of the analysis of this research. In addition, this paper also used financial ratios, other non-financial ratios, and factor analysis to extract adaptable variables. Moreover, the artificial neural network (ANN) and data mining (DM) techniques were used to construct the financial distress prediction model. The empirical experiment with a total of 37 ratios and 68 listed companies as the initial samples obtained a satisfactory result, which testifies for the feasibility and validity of our proposed methods for the financial distress prediction of listed companies.

This paper makes four critical contributions: (1) The more factor analysis we used, the less accuracy we obtained by the ANN and DM approach. (2) The closer we get to the actual occurrence of financial distress, the higher the accuracy we obtain, with an 82.14% correct percentage for two seasons prior to the occurrence of financial distress. (3) Our empirical results show that factor analysis increases the error of classifying companies that are in a financial crisis as normal companies. (4) By developing a financial distress prediction model, the ANN approach obtains better prediction accuracy than the DM clustering approach. Therefore, this paper proposes that the artificial intelligent (AI) approach could be a more suitable methodology than traditional statistics for predicting the potential financial distress of a company.

1. Introduction

In Taiwan, domestic and foreign capital markets have developed rapidly in recent years, gradually giving people the idea of making a financial investment. There are various financial investment objects, such as stocks, futures, options, bond funds etc., and investment stock is the most widely accepted in society. However, capital markets are volatile, and most investors only know that a company is in financial trouble after the financial statement of the company has been made public. Therefore, forecasting corporate financial distress plays an increasingly important role in today's society since it has a significant impact on lending decisions and the profitability of financial institutions. The ability to make accurate bankruptcy predictions are of critical importance to various professionals, such as bank loan officers, creditors, stockholders, bondholders, financial analysts, governmental officials, as well as the general public, as it provides them with timely warnings (Ko & Lin, 2006).

Financial failure occurs when a firm suffers chronic and serious losses or when the firm becomes insolvent with liabilities that are disproportionate to its assets (Hua, Wang, Xu, Zhang, & Liang, 2007). Common causes and symptoms of financial failure include lack of financial knowledge, failure to set capital plans, poor debt management, inadequate protection against unforeseen events and difficulties in adhering to proper operating discipline in the financial market. The common assumption underlying bankruptcy prediction is that a firm's financial statements appropriately reflect above characteristics. Several classification techniques have been suggested to predict financial distress using ratios and data originating from these financial statements, e.g., univariate approaches (Beaver, 1966), multivariate approaches, linear multiple

* Corresponding author. Tel.: +886 3 5820100; fax: +886 3 5610616.
E-mail address: wschen@itri.org.tw (W.-S. Chen).

0957-4174/$ - see front matter
doi:10.1016/j.eswa.2008.03.020

discriminant approaches (MDA) (Altman, 1968; Altman, Edward, Haldeman, & Narayanan, 1977), multiple regression (Meyer & Pifer, 1970), logistic regression (Dimitras, Zanakis, & Zopounidis, 1996), factor analysis (Blum, 1974), and stepwise (Laitinen & Laitinen, 2000). However, strict assumptions of traditional statistics such as linearity, normality, independence among predictor variables and pre-existing functional form relating to the criterion variable and the predictor variable limit their application in the real world (Hua et al., 2007).

With radical changes taking place in corporate finance and the global economic environment, critical financial ratios can change dynamically (John & Robert, 2001). This means that it is both important as well as necessary to develop an evolutionary approach for coping with future dynamic financial environments. Therefore, this paper proposes a model of financial distress prediction integrating artificial neural network (ANN) and data mining (DM) techniques. The main objectives of this paper are to (1) adopt ANN and DM techniques to construct a financial distress prediction model, (2) use financial and non-financial ratios to enhance the accuracy of the financial distress prediction model, (3) employ a traditional statistical method (factor analysis) to compare the degree of accuracy with that of the artificial intelligent (AI) approach, and (4) to expand this model so that it will work within a financial distress prediction system to provide information to investors as well as investment monitoring organizations. The data for our experiment were collected from the Taiwan stock exchange corporation (TSEC) database.

The rest of this paper is organized as follows. A literature review of related studies is provided in Section 2. Section 3 describes our proposed approach and the functionalities of each process. Section 4 presents the process for selecting suitable indicators by factor analysis. To prove the prediction performance of our approach, we carried out several experiments which are described in Section 5. In Section 6, we compared our results with the ANN, and DM approaches. Finally, in Section 7 we draw our conclusions about financial distress forecasting and discuss future work.

2. Literature review

2.1. Artificial neural network

The ANN is composed of richly interconnected non-linear nodes that communicate in parallel. The connection weights are modifiable, allowing ANN to learn directly from examples without requiring or providing an analytical solution to the problem. The most popular forms of learning are:

- *Supervised learning*: Patterns for which both their inputs and outputs are known are presented to the ANN. The task of the supervised learner is to predict the value of the function for any valid input object after having seen a number of training examples. ANN employing supervised learning has been widely utilized for the solution of function approximation and classification problems.
- *Unsupervised learning*: Patterns are presented to the ANN in the form of feature values. It is distinguished from supervised learning by the fact that there is no *a* priori output. ANN employing unsupervised learning has been successfully employed for data mining and classification tasks. The self-organizing map (SOM) and adaptive resonance theory (ART) constitutes the most popular exemplar of this class.

A back-propagation network (BPN) is a neural network that uses a supervised learning method and feed-forward architecture. A BPN is one of the most frequently utilized neural network techniques for classification and prediction (Wu, Yang, & Liang, 2006), and is considered an advanced multiple regression analysis that can accommodate complex and non-linear data relationships (Jost, 1993). It was first described by Werbos (1974), and further developed by Ronald, Rumelhart, and Hinton (1986). The details for the back-propagation learning algorithm can be found in Medsker and Liebowitz (1994).

Fig. 1 shows the $l-m-n$ (l denotes input neurons, $m$denotes hidden neurons, and n denotes output neurons) architecture of a BPN model (Panda, Chakraborty, & Pal, 2007). The input layer can be considered the model stimuli and the output layer the input stimuli outcome. The hidden layer determines the mapping relationships between input and output layers, whereas the relationships between neurons are stored as weights of the connecting links. The input signals are modified by the interconnection weight, known as weight factor w_{ji}, which represents the interconnection of the ith node of the first layer to the jth node of the second layer. The sum of the modified signals (total activation) is then modified by a sigmoid transfer function (f). Similarly, the output signals of the hidden layer are modified by interconnection weight w_{kj} of the kth node of the output layer to the j th node of the hidden layer. The sum of the modified signals is then modified by sigmoid transfer (f) function and the output is collected at the output layer.

Let $I_p = (I_{p1}, I_{p2}, \ldots, I_{pl})$, $p = 1,2,\ldots,N$ be the pth pattern among N input patterns. Where w_{ji} and w_{kj} are connection weights between the ith input neuron to the jth hidden neuron, and the jth hidden neuron to the kth output neuron, respectively (Panda et al., 2007).

Output from a neuron in the input layer is

$$O_{pi} = I_{pi}, \quad i = 1,2,\ldots,l \tag{1}$$

Output from a neuron in the hidden layer is

$$O_{pj} = f(NET_{pj}) = f\left(\sum_{i=0}^{1} w_{ji} o_{pi}\right), \quad j = 1,2,\ldots,m \tag{2}$$

Output from a neuron in the output layer is

$$O_{pk} = f(NET_{pk}) = f\left(\sum_{j=0}^{m} w_{kj} o_{pj}\right), \quad k = 1,2,\ldots,n \tag{3}$$

Where $f(\,)$ is the sigmoid transfer function given by $f(x) = 1/(1 + e^{-x})$.

BPN has been applied to various areas, such as investigating long-term tidal predictions (Lee, 2004), improving customer satisfaction (Deng, Chen, & Pei, 2007), predicting flank wear in drills (Panda et al., 2007), enhancing job completion time prediction in the semiconductor fabrication factory (Chen, 2007), and providing

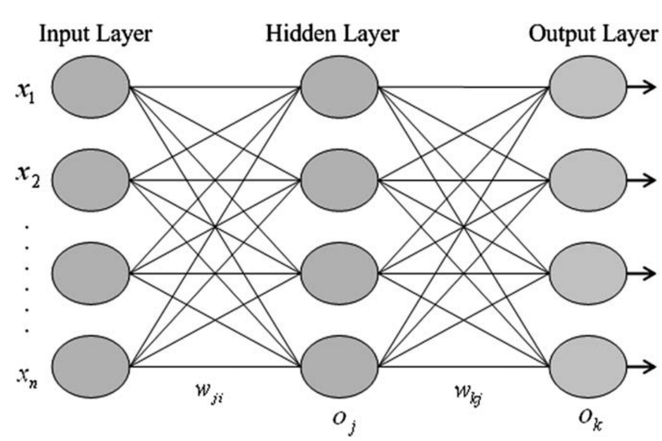

Fig. 1. Back-propagation network architecture.