



Innovative Applications of O.R.

# Prediction of financial distress: An empirical study of listed Chinese companies using data mining

Ruibin Geng<sup>a</sup>, Indranil Bose<sup>b,1</sup>, Xi Chen<sup>a,\*</sup><sup>a</sup> School of Management, Zhejiang University, 866 Yuhangtang Road, Hangzhou, Zhejiang Province 310058, People's Republic of China<sup>b</sup> Indian Institute of Management Calcutta, Diamond Harbour Road, Joka, Kolkata 700104, West Bengal, India

## ARTICLE INFO

## Article history:

Received 1 August 2013

Accepted 10 August 2014

Available online 15 August 2014

## Keywords:

Chinese companies

Financial distress

Financial indicators

Neural network

Majority voting

## ABSTRACT

The deterioration in profitability of listed companies not only threatens the interests of the enterprise and internal staff, but also makes investors face significant financial loss. It is important to establish an effective early warning system for prediction of financial crisis for better corporate governance. This paper studies the phenomenon of financial distress for 107 Chinese companies that received the label 'special treatment' from 2001 to 2008 by the Shanghai Stock Exchange and the Shenzhen Stock Exchange. We use data mining techniques to build financial distress warning models based on 31 financial indicators and three different time windows by comparing these 107 firms to a control group of firms. We observe that the performance of neural networks is more accurate than other classifiers, such as decision trees and support vector machines, as well as an ensemble of multiple classifiers combined using majority voting. An important contribution of the paper is to discover that financial indicators, such as net profit margin of total assets, return on total assets, earnings per share, and cash flow per share, play an important role in prediction of deterioration in profitability. This paper provides a suitable method for prediction of financial distress for listed companies in China.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

Prediction of financial distress has been a topic of interest over the decades because of its great importance to listed companies, interested stakeholders and even the economy of a country (Wanke, Barros, & Faria, 2014). If the prediction of financial distress is reliable, managers of firms can initiate remedial measures to avoid deterioration before the crisis, and investors can grasp the profitability situation of the listed companies and adjust their investment strategies to reduce anticipated investment related losses. However, the rapid development of the capital market and the integration of the global economy have increased the number of companies that suffer from financial distress over the years. In October 2007, the stock market in China crashed and wiped out more than two-thirds of its market value. According to the NASDAQ Company List, among 578 listed companies in 2008, the number of loss-making enterprises reached 278. Besides, Begbies Traynor's latest Red Flag Alert shows a 24% increase in the UK companies facing 'critical' levels of financial distress in the last

quarter of 2011 compared to the last quarter of 2010. It is not a surprise that numerous companies have faced consecutive years of loss, business damage, interests and assets shrinking, suspension of listing every year from 2000 to 2011. With the emergence of China as one of the leading markets for international investors, financial distress of Chinese companies has attracted increasing attention. Therefore, discovery of a suitable model for predicting the financial distress of listed Chinese companies is likely to be of great significance to global investors.

Financial distress of a company usually refers to the situation that operating cash flow of a company cannot supersede the negative net assets of the firm. Different countries have different accounting procedures and rules, and the definition of financial distress put forward by different scholars is not always the same. It is generally agreed on that financial failure leads to substantive weakening of profitability of the company over time. Bankruptcy is the most widely used outcome of financial distress of a company. Ball and Foster (1982) pointed out that taking bankruptcy as a criterion ignored other options that a firm had, such as reducing its scale of operations, liquidating all of its assets and seeking a merger, if it faced "long-term cash flow problems". China Securities Regulatory Commission carries out a 'Special Treatment' (ST) warning mechanism to indicate abnormalities in a listed companies' financial status. The aim of ST is to warn the managers

\* Corresponding author. Tel.: +86 571 88206827.

E-mail addresses: [grace.bin1207@gmail.com](mailto:grace.bin1207@gmail.com) (R. Geng), [indranil\\_bose@yahoo.com](mailto:indranil_bose@yahoo.com) (I. Bose), [chen\\_xi@zju.edu.cn](mailto:chen_xi@zju.edu.cn) (X. Chen).<sup>1</sup> Tel.: +91 33 2467 8300.

and investors and also to use external mechanism to force the firm to enhance its operating performance.

Empirical studies using bankruptcy as the outcome of financial distress has occupied a major part of literature related to this area. However, scholars in China have found it difficult to get enough bankruptcy related data due to the late implementation and incompleteness of the Chinese bankruptcy law. It is believed that ST regulation in the China's stock market provided a unique opportunity to provide early warning of impending financial crisis to be faced by company. According to the Chinese Stock Listing Exchange Rule, there are three main reasons for designation of a company as ST: (1) a listed company has negative net profits for 2 years consecutively; (2) the shareholders' equity of the company is lower than the registered capital; and (3) a firm's operations have stopped and there is no hope of restoring operations in the next 3 months due to natural disasters, serious accidents, or lawsuits and arbitration. There are three types of ST labels to represent different levels of financial distress, that is, ST, \*ST, and S\*ST. ST means special treatment due to loss for two consecutive years. \*ST represents delisting warning due to loss for three consecutive years. S\*ST indicates \*ST stocks which are yet to restructure share rights. According to the Chinese regulation, if a ST firm cannot improve its performance within the next 3 years, it is labeled as PT (Particular Transfer) and may be delisted from the stock market. Once the companies are delisted, the stockholders suffer severe losses. So identifying firms that are likely to be labeled ST in advance is an important activity for investors.

The connection between financial distress and bad profitability has been studied in extant research on Chinese firms. Ding, Song, and Zen (2008) discovered that Chinese ST companies that showed the sign of deterioration in the year before receiving such a label had a high probability of falling into financial distress in the following year. Prediction of ST could provide an earlier warning about bankruptcy by reminding the firms to pay attention to their profitability. So we chose to use ST companies as examples of financially distressed firms who could move on to become bankrupt in future. The objective of this paper is to apply data mining models to forecast which companies are likely to receive the ST label. Such models can be useful in helping investors manage risks and make decisions in a more informed way. We aim to answer three questions in this paper: which data mining models can predict financial distress of Chinese firms most accurately; which financial indicators are most effective in this prediction; and how early can we foresee the sign of deterioration of profitability of the company? The results of this research can provide an early warning to listed companies so that they can take corrective actions in order to avoid financial distress and delisting.

The reminder of this paper is organized as follow. In Section 2 related studies on prediction of financial distress using data mining techniques are reviewed. The research framework and empirical data collected for ST prediction are described in Section 3. Results of the numerical experiments are described in Section 4. Finally, discussion of the empirical results and their implications are presented in Sections 5 and 6.

## 2. Literature review

Financial distress is a broad concept that contains several situations in which firms face financial difficulty. The most common terms used to describe these situations are 'bankruptcy', 'failure', 'insolvency', and 'default'. Altman (1993) put forward a complete description and definition of financial distress and pointed out that bankruptcy was closest to the legal definition of financial distress. Zmijewski (1984) defined financial distress as the act of filing a petition for bankruptcy. However, many financially distressed

firms never filed for bankruptcy, due to acquisition or privatization, whereas healthy firms often filed for bankruptcy to avoid taxes and expensive lawsuits (Theodossiou, Kahya, Saidi, & Philippatos, 1996). On the other hand, 'failure' was a situation where a firm could not pay lenders, preferred stock shareholders, suppliers, etc., or a bill was overdrawn, or the firm was bankrupt. All these situations resulted in a discontinuity of the firm's operations (Dimitras, Zanakis, & Zopounidis, 1996). However, Altman (1993) defined 'failure' as the situation where "the realized rate of return on invested capital, with allowances for risk consideration, is significantly and continually lower than prevailing rates of similar investments", which did not indicate the discontinuity of a firm. For example, when Washington Mutual failed in 2008, the FDIC was able to broker a deal in which JP Morgan Chase bought the assets of Washington Mutual for US\$ 1.9 billion (Palmeri, 2008). 'failure Insolvency' represented negative performance indicating problems related to liquidity and was synonymous with negative net worth (Zopounidis & Dimitras, 1998). Entities most commonly became insolvent by taking on too much debt. One example of insolvency was Almeico Ltd. that failed to pay Fimbank a debt that was due in 2013 (Orr, 2003). 'Default' referred to a situation where a firm violated the condition of an agreement with a creditor that resulted in legal action. For example, the Ministry of Corporate Affairs in India published a list of companies that defaulted in the past (Ministry of Corporate Affairs, 2014). Ross, Westerfield, and Jaffe (2002) stated that "financial distress is a situation where a firm's operating cash flows were not sufficient to satisfy current obligation (such as trade credit or interest expenses) and the firm is forced to take corrective action". To summarize, 'default' essentially meant a debtor had not paid a debt which he or she was required to have paid. 'Insolvency' was a legal term meaning that a debtor was unable to pay his or her debts. 'Bankruptcy' was a legal finding that imposed court supervision over the financial affairs of those firms that were insolvent or were in default.

ST label has been used as the symbol of financial distress in several studies related to Chinese firms (Altman, Heine, Zhang, & Yen, 2007; Bailey, Huang, & Yang, 2011; Sun & Li, 2008a). ST firms go through one or more of four stages that include omission or reduction of the annual dividend payment due to cash shortage, default on loan payments leading to a lawsuit, reorganization or take over and deletion from a stock exchange, and transfer to Asset Management Companies for disposal (Altman et al., 2007). There are two main reasons for taking ST companies as examples as financially distressed firms. Firstly, the financial deterioration of a company is generally considered to be a gradual process. ST as an *ex ante* event of bankruptcy is a good representation of financial distress. In fact, the symptoms of Chinese ST companies are close to the definition of financial distress given by Newton (1975), Gestel et al. (2006), and Lau (1987). Secondly, researchers in China find it difficult to obtain data related to bankrupt firms due to the process of bankruptcy in China. Dairui and Jia (2009) define financially distressed companies as ST and PT companies because they lacked a database of bankruptcy related information on Chinese listed companies.

Statistical techniques have been commonly used for prediction of business failure. Discriminant analysis has been the most frequently used method before 1980s (Altman, 1968; Beaver, 1966). This method has been criticized for its unrealistic assumptions, such as linear separability, multivariate normality, and independent predictor variables, which did not hold in case of real applications. At the same time this method did not provide any estimate of the associated risk of failure. To overcome such limitations, researchers have proposed linear conditional probability models (LPM) and logit or probit regression analysis. Meyer and Pifer (1970) employed LPM for prediction of bankruptcy. Logit analysis was proposed by Martin (1977) for the prediction of bank failures,

and for prediction of business failures by Ohlson (1980). Dimitras et al. (1996) provided a thorough review of statistical methods used for prediction of business failures. Logistic regression is widely used to predict the probability of financial distress (Li, Crook, & Andreeva, 2013).

Research methods for prediction of financial distress advanced from univariate to multivariate models using machine learning methods based on artificial intelligence. According to Ravi Kumar and Ravi (2007), researchers employed different types of intelligent techniques to solve this problem (Carlos, 1996; Fletcher & Goss, 1993; Wilson & Sharda, 1994) and neural network (NN) was the most commonly used technique. Other data mining techniques included decision trees (DT) (Frydman, Altman, & Kao, 1985), case based reasoning (Jo, Han, & Lee, 1997; Li & Sun, 2009), genetic algorithms (Shin & Lee, 2002), Kohonen map (du Jardin & Séverin, 2012), rough sets (McKee, 2000), simulation analysis (Cohen, Doumpos, Neofytou, & Zopounidis, 2012), and support vector machines (SVM) (Fan & Palaniswami, 2000; Gestel et al., 2006). Among them, the rough sets approach is distinctive as it is based on rough sets theory rather than statistical probability or fuzzy set theory. There are two analysis methods based on rough sets theory that have been applied successfully for prediction of bankruptcy (Greco, Matarazzo, Slowinski, & Zanakakis, 2011). These include the Dominance-based Rough Set Approach (DRSA) and the Variable Consistency Dominance-based Rough Set Approach (VC-DRSA). These approaches are useful when monotonic relationships exist between condition attributes and the decision classes. The rough approximations defined within DRSA (Błaszczyński, Greco, & Słowiński, 2007) require that objects having not-worse evaluation with respect to a set of condition attributes than a referent object cannot be assigned to a worse class than the referent object. However, the strict application of dominance principle by DRSA may identify several cases as unclassified. The problem becomes severe particularly when the data sets used to generate dominance principle are large. VC-DRSA extended DRSA by relaxing the strict dominance principle and admitting some inconsistent objects into the lower approximations (Greco, Matarazzo, Slowinski, & Stefanowski, 2001).

NN is inspired by the biological networks of the human nervous system. As a nonlinear mathematical approach, NN often outperforms other single classifiers when testing complex data patterns as evidenced by Bose and Pal (2006) and Chen and Du (2009). Dhar, Mukherjee, and Ghoshal (2010) pointed out that the optimization parameter selection in the model building process, such as changing the number of neurons for each layer, the number of hidden layers, and the learning rate, made the performance better than comparable methods. Methods for combination of multiple classifiers have been regarded as a new direction to develop highly reliable decision making systems (Tax, Van Breukelen, Duin, & Kittler, 2000; Tsai, 2014). Existing combination methods included majority voting (MV), Bayesian techniques, particle swarm optimization, and SVM (Shie, Chen, & Liu, 2012). Extant research has indicated that the performance of individual classifiers will be improved significantly by combining the multiple classifiers in the area of financial distress prediction (Sun & Li, 2008b; Xiao, Yang, Pang, & Dang, 2012). The main reason arises from two different aspects (Xu, Krzyzak, & Suen, 1992). First, there are a number of classification algorithms available for financial distress prediction. These algorithms are based on different theories and methodologies. When classifiers are used on identical data set, the classifiers estimate the same class posterior probability. To suppress the errors in estimation, the classifier outputs can be combined. Second, companies' financial situations can be described using numerous types of features, such as continuous variables, binary values, and discrete labels. One single classifier does not perform well for all types of representations. DT is more suitable for decision making based

on nominal data. The rough sets approach requires monotonicity of criteria. As a result, different classifiers are needed to handle different types of features. Table 1 summarizes the representative research papers that have used single or multiple classifiers for prediction of financial distress for Chinese firms.

Existing studies on prediction of financial distress have suffered from a few limitations. The first issue has been limited availability of data on financial distressed companies. For example, in the case of Chinese listed companies, ST companies were the best choice for financial distress analysis. However, only dozens of companies were designated as ST every year. The number of ST companies operating in the stock market was much smaller than that of non-ST companies. The second issue has been related to the selection of indicators of financial distress. Extant research has primarily focused on discovery of 'better' models for prediction of financial distress without paying much attention to the analysis of the predictive ability of different financial indicators. Feature selection methods adopted in previous financial distress prediction studies are either expert recommendation or statistical methods such as principle component analysis (Lin, Liang, Yeh, & Huang, 2014). Thirdly, existing research papers mostly collected financial data of ST companies that were either 1 year before treatment or 2 year before treatment for prediction of financial distress (Ding et al., 2008). However, it was not very meaningful to forecast ST risk 1 or 2 years ahead, because ST was determined by the financial status of the company in the previous 2 years. Sevim, Oztekin, Bali, Gumus, and Guresen (2014) collected Turkish financial distress cases covering the period of 1992–2011 but achieved only a 12 month warning system. There was a need to collect more historical data and make earlier predictions using those data. Finally, a rigorous selection of the control sample was almost absent in past research. A common practice has been to select control samples of similar size by using companies belonging to the same industrial sector.

### 3. Research methodology

In this paper, we conducted the prediction of financial distress using the process of knowledge discovery. NN, DT and SVM are selected as the best three single classifiers for this purpose. Multiple classifiers based on MV (Sun & Li, 2008b) are introduced to combine the prediction results obtained from the best three models. We evaluated the performance of these four methods based on four different partition ratios (i.e., 60:40, 70:30, 80:20, and 90:10). In order to enhance the robustness of the results and limit overfitting, we used 10 times repeated random sub-sampling validation. An important objective of this research is to determine whether we could foresee the sign of financial deterioration of the firm early on? So we collected financial data of the listed companies 3 years, 4 years, and 5 years prior to the year when the companies received ST and built prediction models for each of these time periods separately. We also measured the predictive ability of the variables using importance ranking and identified the most important indicators of financial distress. Fig. 1 is a schematic diagram showing the steps followed in this research.

#### 3.1. Data collection

##### 3.1.1. Sample of financially distressed companies

We selected 107 listed companies from CSMAR (China Security Market Accounting Research) database that were labeled as ST or \*ST during 2008–2011 for reasons such as 2 years of losses, damaged business, or financially bankrupt (excluding major litigation or other factors). The number of companies labeled as ST or \*ST

**Table 1**

Summary of studies on prediction of financial distress for companies.

Papers	Techniques								Sample selection	Best model
	NN	SVM	DT	LR	DA	RSA	CM	MC		
Lin et al. (2014)	Yes	Yes	No	Yes	Yes	No	No	No	Distressed companies with any of following conditions: full-value delivery, stock transaction suspension, re-construction, bankruptcy or withdrawal from stock market	Models based on feature sets selected by an integrated feature selection approach
Chen (2013)	No	No	No	No	No	No	Yes	No	Electronics companies listed on Taiwan stock exchange	Subtractive clustering using particle swarm optimization and fuzzy systems
Xiao et al. (2012)	Yes	Yes	No	Yes	No	Yes	No	Yes	ST Chinese listed companies	Multiple classifiers based on Dempster–Shafer evidence theory
Li and Sun (2009)	No	No	Yes	Yes	Yes	No	No	Yes	ST Chinese listed companies	Hybrid case based reasoning system
Chen and Du (2009)	Yes	No	No	No	No	No	Yes	No	Companies stopped and suspended according to Taiwan stock exchange rules	Neural networks
Ding et al. (2008)	Yes	Yes	Yes	Yes	Yes	No	No	No	ST Chinese listed companies	Support vector machines
Sun and Li (2008b)	Yes	Yes	Yes	Yes	Yes	No	No	Yes	ST Chinese listed companies	Multiple classifiers based on majority voting
Bose and Pal (2006)	Yes	Yes	No	No	Yes	No	No	No	Survived and failed click-and-mortar companies	Neural networks
Ahn, Cho, and Kim (2000)	Yes	No	No	No	No	Yes	No	No	Korean firms selected from Korea Information Service	Hybrid model integrating rough sets and neural networks

Neural networks (NN), Support vector machines (SVM), Decision tree (DT), Logistic regression (LR), Discriminant analysis (DA), Rough Set Analysis (RSA), Clustering methods (CM), Multiple classifiers (MC).

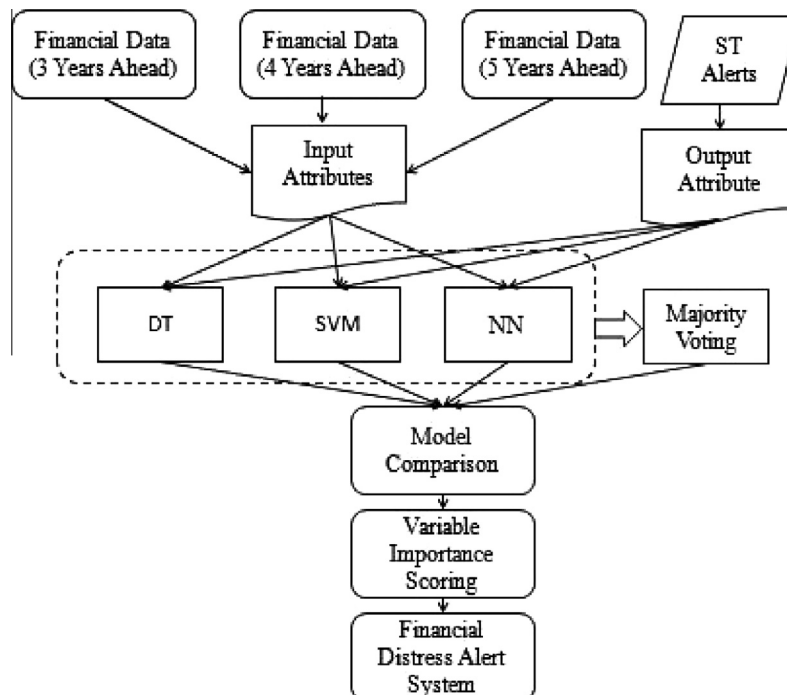
in each year from 2008 to 2011 were 26, 29, 38, and 14. More details about these companies appear in [Appendix A](#).

### 3.1.2. Sample of financial healthy companies

To create a control sample with which we could compare the performance of the financially distressed companies, we selected 107 listed companies randomly from the CSMAR database. These companies had not been labeled as ST till 2011. More details about these companies appear in [Appendix B](#).

### 3.1.3. Time span of the data set

Previous studies on ST Chinese companies could achieve high prediction accuracy by using financial data that was obtained 1 or 2 years prior to the year when the companies received the ST label (Ding et al., 2008), because whether a company was to receive the ST label or not was determined exactly by its financial status in the previous 2 years. We implemented the models based on financial data obtained 3, 4, and 5 years before the companies received the ST label. We used time windows to collect the



**Fig. 1.** Schematic diagram showing the steps of the research.



financial data of ST companies. A 5-year time window meant that we used the financial data of the year  $t - 5$  to predict whether the firm was labeled as ST in year  $t$ . For example, if the company received the ST label in 2008, its financial data in 2003 was used. The data sets based on 3-year and 4-year time windows were collected similarly. As for the control group, the financial data of the financial healthy companies in 2006 (5 years before 2011), in 2007 (4 years before 2011), and in 2008 (3 years before 2011) were collected.

### 3.2. Financial indicators

We chose 31 financial indicators as inputs based on the qualitative analysis of China Accounting Standards. These indicators could be divided into six categories that reflected the company's solvency, profitability, operational capabilities, business development capacity, structural soundness, and capital expansion capacity. A list of the 31 financial indicators is provided in Table 2. All variables used for calculation of the financial indicators are obtained from the balance sheets, income statements or cash flow statements of the companies.

### 3.3. Modeling

The steps of this research are obtained by following the important steps of the standard data mining process of CRISP-DM (Shearer, 2000), including data understanding, data preparation, data modeling, and evaluation.

#### 3.3.1. Data understanding

'Whether a company should receive the ST label or not' is the target variable of prediction. Except this target variable which is a binary variable, all input variables are continuous. Some

descriptive statistics related to the input variables for the three different time windows are shown in Appendix C.

#### 3.3.2. Data preparation

The data records of ST companies and normal companies are combined as the initial data set. The completeness of the data is over 95% and the records with missing data are discarded.

As per the needs of supervised learning, we randomly partitioned the data into two parts for training and testing. The training data is used for building the learning models whereas the testing data is used for testing the predictive ability of the models. Different training to testing ratios such as 60:40, 70:30, 80:20, and 90:10 are used in order to avoid the limitations of undertraining or overtraining. In addition, we created 10 random training samples and 10 random testing samples based on each partition ratio to conduct cross-validation. We repeated all experiments on these 10 paired training and testing samples separately and sequentially. Then we determined the average results across all samples. This type of resampling technique is known as repeated random subsampling validation and has been used in extant research (Usai, Goddard, & Hayes, 2009).

#### 3.3.3. Data modeling

As the first step of data modeling, 10 models based on statistical probability theory, namely, NN, C5.0 DT, C&R Tree, QUEST, CHAID, logistic regression, decision list, Bayes net, discriminant analysis, and SVM are applied for prediction of financial distress. Besides, VC-DRSA and DRSA are also adopted as the representative algorithms based on rough sets theory. We used the jMAF software to generate the results using DRSA and VC-DRSA. NN based on the prune algorithm, C5.0, and SVM outperformed the other techniques and were chosen as representatives of data mining models based on statistical probability. In order to combine the results of

**Table 2**  
List of financial indicators.

Types	Symbols	Formulae for calculation
Solvency	TL/TA	Total liabilities/total assets
	CA/CL	Current assets/current liabilities
	(CA-I)/CL	(Current assets-inventory)/current liabilities
	TL/TSE	Total liabilities/total shareholders' equity
	CL/TA	Current liabilities/total assets
	NOCF/CL	Net operating cash flow/current liabilities
Profitability	EBIT/IE	Earnings before interest and tax (EBIT)/interest expense
	(SR-SC)/SR	(Sales revenue-sales cost)/sales revenue
	NP/SR	Net profit/sales revenue
	EBIT/ATA	Earnings before income tax/average total assets
	NP/ATA	Net profit/average total assets
	NP/ACA	Net profit/average current assets
	NP/AFA	Net profit/average fixed assets
Operational capabilities	NP/ASE	Net profit/average shareholders' equity
	MBI/ATA	Main business income/average total assets
	SR/ACA	Sales revenue/average current assets
	SR/AFA	Sales revenue/average fixed assets
	MBC/AI	Main business cost/average inventory
	MBI/ABAR	Main business income/average balance of accounts receivable
Business development capacity	CS/APA	Cost of sales/average payable accounts
	MBI(t)/MBI( $t - 1$ )	Main business income of this year/main business income of last year
	TA(t)/TA( $t - 1$ )	Total assets of this year/total assets of last year
	NP(t)/NP( $t - 1$ )	Net profit of this year/net profit of last year
Structural soundness	CA/TA	Current assets total assets
	FA/TA	Fixed assets/total assets
	SE/FA	Shareholders' equity/fixed assets
	CL/TL	Current liabilities/total liabilities
Capital expansion capacity	NP/NOS	Net profit/number of ordinary shares at the end of year
	NA/NOS	Net assets/number of ordinary shares at the end of year
	NICCE/NOS	Net increase in cash and cash equivalents/number of ordinary shares at the end of year
	CR/NOS	Capital reserves/number of ordinary shares at the end of year

the top three models, we adopted the process of MV (Sun & Li, 2008b).

### 3.3.4. Evaluation

We measured the performance of the models in terms of their accuracy (the ratio of number of correctly predicted records including ST and non-ST to total number of records), recall (the ratio of number of correctly predicted ST records to total number of actual ST records), and precision (the ratio of number of correctly predicted ST records to total number of predicted ST records), according to the common machine learning evaluation metrics proposed by Davis and Goadrich (2006). The random division of the training and test data might have some impact on the

analysis. So we conducted 10 times repeated random sub-sampling validation based on the same data set. The average predictive performance of the 10 trials is regarded as final prediction result for each model.

## 4. Results and analysis

### 4.1. Prediction performance of data mining techniques

The prediction performance of the NN, DT, and SVM models, and the result of MV based on these three models are shown in Table 3. As mentioned in the earlier section, three measures of

**Table 3**  
Prediction results of NN, DT, SVM, and MV.

$t - 3$				$t - 4$			$t - 5$		
60:40	NN	Accuracy	0.788 (0.044)	NN	Accuracy	0.760 (0.049)	NN	Accuracy	0.762 (0.037)
		Recall	0.797 (0.059)		Recall	0.787 (0.086)		Recall	0.681 (0.075)
		Precision	0.732 (0.067)		Precision	0.727 (0.057)		Precision	0.793 (0.052)
	DT	Accuracy	0.744 (0.060)	DT	Accuracy	0.705 (0.022)	DT	Accuracy	0.623 (0.046)
		Recall	0.728 (0.080)		Recall	0.698 (0.099)		Recall	0.584 (0.122)
		Precision	0.705 (0.116)		Precision	0.689 (0.058)		Precision	0.621 (0.071)
	SVM	Accuracy	0.711 (0.036)	SVM	Accuracy	0.697 (0.062)	SVM	Accuracy	0.619 (0.056)
		Recall	0.777 (0.053)		Recall	0.681 (0.152)		Recall	0.587 (0.098)
		Precision	0.636 (0.068)		Precision	0.680 (0.079)		Precision	0.605 (0.051)
	MV	Accuracy	0.784 (0.049)	MV	Accuracy	0.761 (0.047)	MV	Accuracy	0.705 (0.035)
		Recall	0.799 (0.049)		Recall	0.783 (0.083)		Recall	0.644 (0.060)
		Precision	0.726 (0.070)		Precision	0.732 (0.060)		Precision	0.712 (0.054)
70:30	NN	Accuracy	0.779 (0.055)	NN	Accuracy	0.787 (0.045)	NN	Accuracy	0.756 (0.078)
		Recall	0.798 (0.036)		Recall	0.809 (0.066)		Recall	0.693 (0.106)
		Precision	0.765 (0.083)		Precision	0.760 (0.070)		Precision	0.785 (0.105)
	DT	Accuracy	0.754 (0.033)	DT	Accuracy	0.715 (0.051)	DT	Accuracy	0.607 (0.103)
		Recall	0.733 (0.108)		Recall	0.741 (0.065)		Recall	0.620 (0.162)
		Precision	0.761 (0.038)		Precision	0.694 (0.086)		Precision	0.601 (0.126)
	SVM	Accuracy	0.741 (0.054)	SVM	Accuracy	0.720 (0.068)	SVM	Accuracy	0.606 (0.079)
		Recall	0.775 (0.073)		Recall	0.719 (0.133)		Recall	0.579 (0.145)
		Precision	0.722 (0.080)		Precision	0.704 (0.079)		Precision	0.607 (0.093)
	MV	Accuracy	0.778 (0.0051)	MV	Accuracy	0.773 (0.054)	MV	Accuracy	0.705 (0.081)
		Recall	0.786 (0.057)		Recall	0.802 (0.064)		Recall	0.674 (0.127)
		Precision	0.766 (0.059)		Precision	0.732 (0.066)		Precision	0.709 (0.103)
80:20	NN	Accuracy	0.742 (0.070)	NN	Accuracy	0.790 (0.064)	NN	Accuracy	0.780 (0.086)
		Recall	0.759 (0.102)		Recall	0.797 (0.093)		Recall	0.719 (0.135)
		Precision	0.686 (0.090)		Precision	0.754 (0.078)		Precision	0.763 (0.079)
	DT	Accuracy	0.730 (0.068)	DT	Accuracy	0.724 (0.057)	DT	Accuracy	0.620 (0.063)
		Recall	0.769 (0.127)		Recall	0.703 (0.128)		Recall	0.660 (0.156)
		Precision	0.678 (0.101)		Precision	0.686 (0.085)		Precision	0.572 (0.104)
	SVM	Accuracy	0.715 (0.064)	SVM	Accuracy	0.718 (0.048)	SVM	Accuracy	0.641 (0.065)
		Recall	0.761 (0.099)		Recall	0.728 (0.120)		Recall	0.621 (0.127)
		Precision	0.645 (0.068)		Precision	0.675 (0.085)		Precision	0.591 (0.091)
	MV	Accuracy	0.745 (0.085)	MV	Accuracy	0.789 (0.058)	MV	Accuracy	0.707 (0.049)
		Recall	0.788 (0.095)		Recall	0.795 (0.058)		Recall	0.727 (0.119)
		Precision	0.672 (0.088)		Precision	0.756 (0.084)		Precision	0.647 (0.067)
90:10	NN	Accuracy	0.768 (0.119)	NN	Accuracy	0.778 (0.066)	NN	Accuracy	0.742 (0.073)
		Recall	0.775 (0.129)		Recall	0.824 (0.066)		Recall	0.672 (0.127)
		Precision	0.760 (0.147)		Precision	0.707 (0.113)		Precision	0.821 (0.112)
	DT	Accuracy	0.724 (0.136)	DT	Accuracy	0.688 (0.090)	DT	Accuracy	0.597 (0.117)
		Recall	0.754 (0.174)		Recall	0.723 (0.208)		Recall	0.597 (0.112)
		Precision	0.734 (0.151)		Precision	0.610 (0.198)		Precision	0.653 (0.233)
	SVM	Accuracy	0.696 (0.107)	SVM	Accuracy	0.736 (0.100)	SVM	Accuracy	0.621 (0.138)
		Recall	0.742 (0.179)		Recall	0.823 (0.124)		Recall	0.578 (0.109)
		Precision	0.667 (0.141)		Precision	0.657 (0.153)		Precision	0.694 (0.225)
	MV	Accuracy	0.762 (0.125)	MV	Accuracy	0.748 (0.081)	MV	Accuracy	0.669 (0.091)
		Recall	0.803 (0.127)		Recall	0.834 (0.084)		Recall	0.635 (0.109)
		Precision	0.746 (0.151)		Precision	0.654 (0.153)		Precision	0.737 (0.161)

The results above represent the mean and standard deviation (in brackets) obtained using 10 times repeated random sub-sampling.

prediction performance are reported in this table and they are accuracy, recall, and precision. Scenarios based on different training to testing ratios as well as different time periods are shown as well. As we have made 10 repeated experiments based on each partition ratio, we will get 40 samples to represent the prediction results of the proposed models. In order to determine the best performing technique, we not only compared the average prediction accuracy of the four models but also conducted two-sample *t* test between the accuracy of NN and that of the other four models. The null hypothesis of each *t* test was that the predictive accuracy of NN was higher than that of SVM, DT, MV and RSA. The statistical results indicated that the prediction accuracy of NN was significantly higher than SVM, DT, and RSA in time windows  $t - 3$ ,  $t - 4$ , and  $t - 5$ . However, the null hypothesis is not rejected when comparing the accuracy of NN with MV in the 3-year and 4-year time windows. This indicated that MV performed as well as NN in these experiments.

The average prediction accuracy of NN is highest in each scenario, except for the case of 60:40 in the 4-year time window and the case of 80:20 in the 3-year time window (Fig. 2). The performance of MV ranked second, close to but not exceeding the accuracy obtained using NN significantly. The standard deviation of accuracy increased dramatically when the training to testing ratio increased to 90:10. The large fluctuation of accuracy indicated the existence of overtraining problem in the case of 90:10. In the 3-year time window, the best result (78.82%) is obtained when the training to testing ratio was 60:40. The highest prediction accuracy of the 4-year time window and the 5-year time window (79.0% and 78.0% respectively) occurred when the training to testing ratio was 80:20. Training to testing ratio affects the prediction accuracy significantly. If most of the data is used for training and a smaller portion of the data is used for testing, data discrepancies between training and testing will be enlarged. The model created using high partition ratio is more likely to overfit to the training set and generally have poor predictive performance. It is observed that a higher training to testing ratio led to a deterioration and higher fluctuation (larger standard deviation) in accuracy, and hence the 60:40 ratio is chosen for further analysis (see Fig. 3).

Using training to testing ratio of 60:40, ST could be predicted with an average accuracy of over 75% for the best performing models. One-way ANOVA with Fisher's LSD procedure is used to test whether the prediction accuracy of each model improved significantly as the time period decreased from 5 years to 3 years. The predictive performance of each model for time periods  $t - 3$ ,  $t - 4$ , and  $t - 5$  showed significant difference except those for NN. It seemed that a warning that was one more year ahead did not sacrifice much accuracy when using NN. However, the accuracy for time period  $t - 3$  was relatively higher than the other two time periods for NN. We concluded that the prediction results based on financial data of the 3-year time window were the most accurate. It was also intuitive that the closer we got to the time of obtaining ST, the more precise prediction we made.

#### 4.2. Rough sets analysis versus statistical probability models

As analysis approaches such as DRSA and VC-DRSA require the order of input attributes corresponds to increasing or decreasing order of preference of a decision maker (McKee, 2000), we examine the monotonicity constraints. We randomly partition the training and testing sets using ratios of 60:40, 70:30, 80:20, and 90:10. Ten times repeated random sub-sampling validation is used for each partition ratio as well. We apply the VC-DRSA method at the partition ratio of 60:40 and set the consistency level at 0.9 (consistency level at 1.0 for standard DRSA analysis), as the strict dominance principles by DRSA render several cases as unclassified. When the training to testing ratio is high, all cases are predicted as

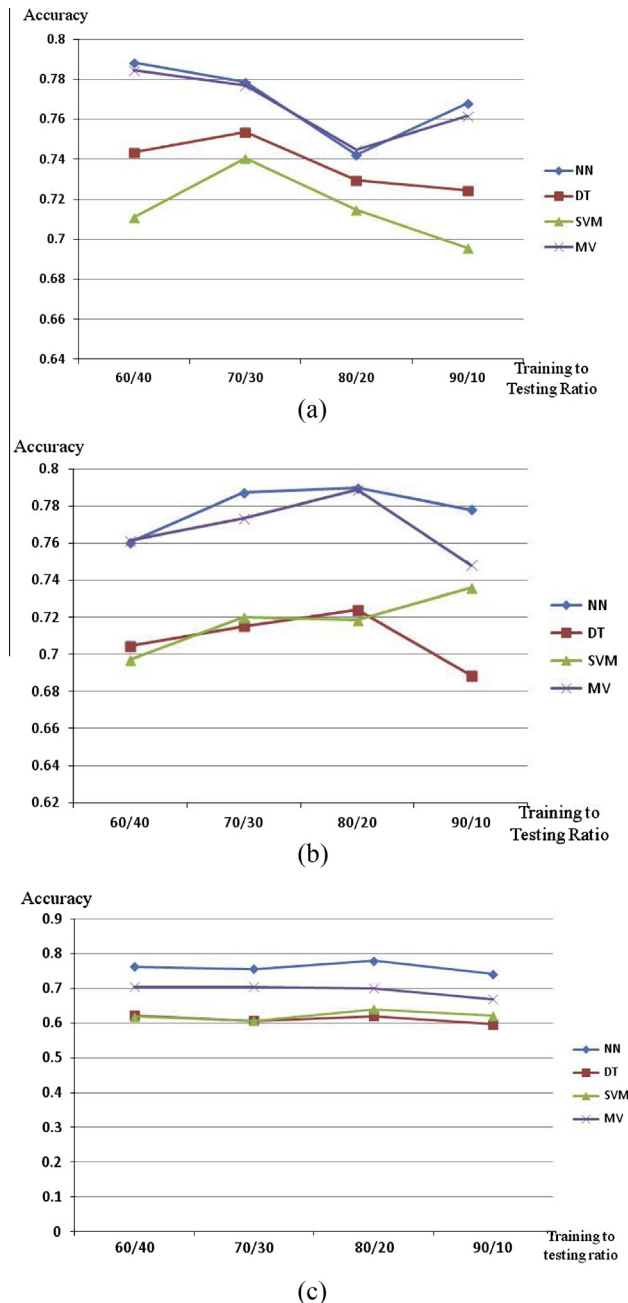


Fig. 2. Predictive accuracy of NN, DT, SVM, and MV for (a) 3-year time window, (b) 4-year time window, and (c) 5-year time window.

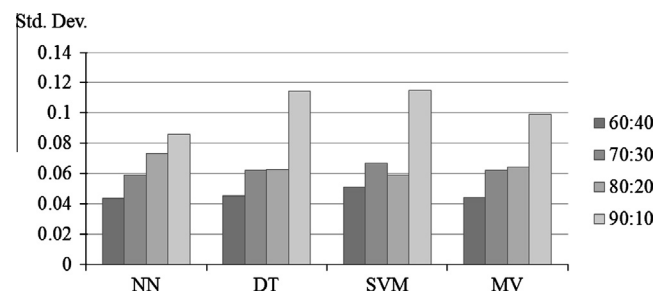


Fig. 3. Average standard deviation of predictive accuracy of NN, DT, SVM, and MV across time windows and for different training to testing ratios.

ST companies by VC-DRSA at any consistency levels, which indicates that VC-DRSA is not a preferred method for this research. Thus we revert to the DRSA method. Prediction results using rough sets analysis for both DRSA and VC-DRSA methods are shown in Table 4. Although the prediction accuracy of DRSA within classified cases is relatively high, nearly a quarter of the cases remain unclassified when the DRSA method is used. Therefore, the total accuracy of rough set analysis is significantly lower than that of SVM, DT, and NN.

#### 4.3. Evaluation of importance of financial indicators

Feature selection is conducted by ranking each financial indicator based on the strength of its relationship to the target variable, independent of other input variables. The idea is to perform a one-way ANOVA  $F$  test for each continuous predictor, which tested if all predictors had the same mean as  $X$ . If  $N_j$  is the number of cases with  $Y = j$ ,  $j = 1 \cdots J$ ,  $\bar{x}_j$  and  $s_j^2$  is the sample mean and sample variance of the predictor  $X$  for target class  $Y = j$ ,  $\bar{x}$  is the grand mean of predictor  $X$ , then the  $F$  statistic is calculated as:

$$F = \frac{\sum_{j=1}^J N_j (\bar{x}_j - \bar{x})^2 / (J - 1)}{\sum_{j=1}^J (N_j - 1) s_j^2 / (N - J)} \quad (1)$$

Predictors are sorted according to their  $p$  value in an ascending order. We applied two ranking methods to evaluate the predictive ability of each financial indicator – score and occurrence. The top seven financial indicators with smallest  $p$  value are assigned scores

of 7 to 1 based on each ranking for 10 repetitions of each model. Financial indicators with higher prediction abilities are synthetically determined by the ranking of final scores and number of times they occurred in the list of most important variables in Table 5. We repeated the process of ranking of financial indicators for the three different time periods. Appendix D in the online supplement shows the results of ranking of financial indicators for NN, DT, SVM, and DRSA. The top ten important indicators generated by the various methods were roughly the same although some differences were observed in the order of their importance.

For the three different time windows, financial indicators with the higher prediction abilities as ranked by the two criteria of score and occurrence are found to be same and included net profit on average total assets (NP/ATA) and return on total assets (EBIT/ATA) (as shown in first two rows of Table 5). Other important financial indicators are mainly from the dimension of capital expansion capacity (e.g., earnings per share, net assets per share, cash flow per share, and capital reserves per share), profitability (e.g., operating gross profit rate, and net profit rate of current assets) and business development capacity (e.g., net profit growth rate).

Based on the ranking obtained for importance of financial indicators, we repeated the accuracy related numerical experiments using the three models after feature selection. For these experiments, only the top ten financial indicators ranked by scores are used as inputs to predict the ST risk, rather than all 31 financial indicators. The training to testing ratio was set as 60:40 to control the problem of overtraining. Predictive performance after feature

**Table 4**  
Prediction results of rough sets analysis based on DRSA and VC-DRSA.

	$t - 3$		$t - 4$		$t - 5$	
	DRSA	VC-DRSA	DRSA	VC-DRSA	DRSA	VC-DRSA
Accuracy	0.689(0.071)	0.751(0.032)	0.589(0.049)	0.757(0.033)	0.538(0.037)	0.692(0.036)
Recall	0.863(0.080)	0.848(0.063)	0.785(0.083)	0.852(0.042)	0.761(0.059)	0.806(0.056)
Precision	0.855(0.084)	0.720(0.061)	0.841(0.041)	0.721(0.057)	0.739(0.078)	0.645(0.055)
Unclassified cases	19.4%(0.077)	0	28.6%(0.044)	0	28.2%(0.044)	0
Accuracy	0.627(0.056)	0.517(0.113)	0.592(0.070)	0.584(0.152)	0.551(0.062)	0.451(0.040)
Recall	0.828(0.055)	1(0)	0.859(0.069)	1(0)	0.744(0.127)	1(0)
Precision	0.792(0.074)	0.513(0.053)	0.816(0.115)	0.580(0.147)	0.703(0.083)	0.451(0.040)
Unclassified cases	23.0%(0.057)	0	28.5%(0.068)	0	25.1%(0.056)	0
Accuracy	0.689(0.071)	0.512(0.084)	0.644(0.092)	0.501(0.070)	0.534(0.076)	0.456(0.065)
Recall	0.863(0.080)	1(0)	0.880(0.069)	1(0)	0.798(0.121)	1(0)
Precision	0.855(0.084)	0.512(0.084)	0.805(0.121)	0.501(0.070)	0.682(0.116)	0.456(0.065)
Unclassified cases	19.4%(0.077)	0	23.2%(0.065)	0	28.0%(0.085)	0
Accuracy	0.637(0.107)	0.498(0.146)	0.630(0.067)	0.526(0.102)	0.513(0.077)	0.471(0.099)
Recall	0.883(0.128)	1(0)	0.800(0.184)	1(0)	0.598(0.130)	1(0)
Precision	0.812(0.187)	0.498(0.146)	0.805(0.121)	0.526(0.102)	0.668(0.166)	0.471(0.099)
Unclassified cases	23.3%(0.119)	0	22.1%(0.066)	0	22.1%(0.077)	0

The results represent the mean and standard deviation (in brackets) obtained using 10 repeated random sub-sampling validation.

**Table 5**  
Importance ranking of financial indicators for time periods  $t - 3$ ,  $t - 4$ , and  $t - 5$ .

Number	$t - 3$		$t - 4$		$t - 5$	
	Ranked by score	Ranked by occurrence	Ranked by score	Ranked by occurrence	Ranked by score	Ranked by occurrence
1	NP/ATA	NP/ATA	EBIT/ATA	EBIT/ATA	EBIT/ATA	EBIT/ATA
2	EBIT/ATA	EBIT/ATA	NP/ATA	NP/ATA	NP/ATA	NICCE/NOS
3	TA(t)/TA( $t - 1$ )	TA(t)/TA( $t - 1$ )	NP/NOS	NICCE/NOS	NICCE/NOS	NP/ATA
4	NP/NOS	NP/NOS	NICCE/NOS	NP/NOS	NP/SR	NA/NOS
5	(SR-SC)/SR	(SR-SC)/SR	CR/NOS	CR/NOS	NA/NOS	(SR-SC)/SR
6	NICCE/NOS	NICCE/NOS	NA/NOS	NA/NOS	(SR-SC)/SR	NP/SR
7	CR/NOS	CR/NOS	NOCF/CL	NP/ACA	MBI(t)/MBI( $t - 1$ )	MBI(t)/MBI( $t - 1$ )
8	NP/SR	NA/NOS	NP/ACA	NOCF/CL	FA/TA	FA/TA
9	NA/NOS	NP(t)/NP( $t - 1$ )	NP(t)/NP( $t - 1$ )	NP(t)/NP( $t - 1$ )	NP/ACA	NP(t)/NP( $t - 1$ )
10	NP(t)/NP( $t - 1$ )	NP/SR	NP/SR	NP/SR	NP(t)/NP( $t - 1$ )	NP/ACA



**Table 6**

Predictive performance of the three models after feature selection.

		NN	Change for NN (%)	DT	Change for DT (%)	SVM	Change for SVM (%)
$t - 3$	Accuracy	77.79% (0.035)	−0.02	74.39% (0.046)	9.74	73.96% (0.041)	6.46
	Recall	78.82% (0.083)	−7.18	76.25% (0.068)	14.24	75.40% (0.087)	12.31
	Precision	75.08% (0.031)	3.93	72.10% (0.074)	10.65	71.06% (0.043)	5.71
$t - 4$	Accuracy	78.45% (0.035)	−0.47	72.87% (0.057)	3.31	70.39% (0.057)	1.04
	Recall	81.82% (0.057)	−0.52	72.98% (0.099)	3.07	70.91% (0.080)	2.25
	Precision	74.64% (0.069)	−0.18	71.45% (0.091)	3.92	67.45% (0.085)	0.28
$t - 5$	Accuracy	74.91% (0.041)	−1.94	68.66% (0.041)	4.43	65.70% (0.049)	9.18
	Recall	69.94% (0.067)	−0.84	65.03% (0.146)	13.14	68.18% (0.125)	6.54
	Precision	73.72% (0.069)	−3.49	66.80% (0.093)	3.10	61.90% (0.094)	12.81

The results represent the mean and standard deviation (in brackets) obtained using 10 times repeated random sub-sampling validation.

selection is shown in Table 6. Deleting the unimportant variables from the prediction models is found to improve the performance of DT and SVM models but reduced the accuracy of the NN model. Paired  $t$  test is conducted to figure out whether there existed significant difference between prediction accuracy before and after feature selection. The  $p$ -value of DT in time periods  $t - 3$  and  $t - 5$ , as well of SVM for time period  $t - 3$  was less than 0.05. The remaining paired  $t$  tests (at the significance level of 0.1) indicated that the prediction accuracy of NN models did not change significantly before and after feature selection. So we can conclude that the performance of NN is not affected by feature selection. This implied that the top ten variables could effectively represent the 31 financial indicators for prediction of financial distress.

## 5. Discussion

In this research we examined three questions: which method performed best in the prediction of financial distress? How early could we detect the financial deterioration of listed companies? Which financial indicators were of importance for determination of financial distress? The results obtained in the previous section indicated that NN was the most robust and efficient model for prediction of financial deterioration, which was in conformity with previous literature (Bose & Pal, 2006; Chen & Du, 2009). In fact, NN outperformed the other data mining methods in the 3-year,

4-year, and 5-year time windows with various training to testing ratios. There were two advantages of NN which might have led to its superior performance. Firstly, NN did not make any assumption about the statistical distribution or properties of the data. Secondly, NN had a strong ability to fit nonlinear data and could approximate complex data patterns accurately. It was likely that the 31 financial indicators had complicated nonlinear relationships with the target variable, 'whether ST or not' and so NN gave rise to good classification due to the presence of complex relationships and noise in the data. SVM has been proposed as an excellent nonlinear relationship recognition method by Vapnik (1995). Studies by Ding et al. (2008) concluded that the performance of SVM in forecasting financial condition of Chinese listed companies was better than that of back-propagation neural networks (BPNN) when training to testing ratio was 50:50. Various training algorithms can be used for NN modeling. We tried different training algorithms, including prune algorithm, multilayer perceptron (MLP), and radial basis function network (RBFN), and determined that the prune algorithm outperformed others. The prune algorithm started with a large network and then removed the weakest units in the hidden and input layers. In our experiments, the best performance was obtained when there were two hidden layers for NN after pruning. However, DT models were able to generate if-then rules that conveyed insightful knowledge. In order to detect robust rules, we filtered the rules that determined high risk

**Table 7**

Robust rules generated by DT.

$t - 3$		$t - 4$		$t - 5$	
Ranked by confidence	Ranked by coverage	Ranked by confidence	Ranked by coverage	Ranked by confidence	Ranked by coverage
If NP/ATA $\leq$ 0.025 and NICCE/NOS $\leq$ 0.007 then ST = TRUE	If NP/ATA $\leq$ 0.025 and NICCE/NOS $\leq$ 0.414 then ST = TRUE	If NOCF/CL $\leq$ 0.217 and NP(t)/NP(t - 1) $\leq$ -0.332 then ST = TRUE	If NP/ACA $\leq$ 0.066 and NP/NOS $\leq$ 0.191 then ST = TRUE	If MBI(t)/MBI(t - 1) $\leq$ 0.186 and NICCE/NOS $\leq$ -0.121 then ST = TRUE	If NICCE/NOS $\leq$ -0.121 then ST = TRUE
If NOCF/CL $\leq$ 0.186 and NP/SR $\leq$ 0.041 and NP/ATA $\leq$ 0.044 then ST = TRUE	If NP/ASE $\leq$ 0.055 then ST = TRUE	If NOCF/CL $\leq$ 0.205 and NP(t)/NP(t - 1) $\leq$ -0.338 then ST = TRUE	If NOCF/CL $\leq$ 0.207 and SR/ACA > 0.350 and NP/NOS $\leq$ 0.092 then ST = TRUE	If NP/AFA $\leq$ 0.017 then ST = TRUE	If EBIT/ATA $\leq$ 0.085 and FA/TA $\leq$ 0.585 and NICCE/NOS $\leq$ 0.051 then ST = TRUE
If NP/ATA $\leq$ 0.025 and CA/TA > 0.366 then ST = TRUE	If NP/AFA $\leq$ 0.067 then ST = TRUE	If NP/ATA $\leq$ 0.025 and CS/APA > 4.043 and FA/TA $\leq$ 0.669 then ST = TRUE	If NP(t)/NP(t - 1) $\leq$ -0.328 then ST = TRUE	If NP/ASE $\leq$ 0.010 then ST = TRUE	If MBI(t)/MBI(t - 1) $\leq$ 0.186 and NICCE/NOS $\leq$ -0.121 then ST = TRUE
If NP/ATA $\leq$ 0.025 and FA/TA $\leq$ 0.632 then ST = TRUE	If NP/ATA $\leq$ 0.025 and FA/TA $\leq$ 0.632 then ST = TRUE	If NOCF/CL $\leq$ 0.207 and CL/TL > 0.929 and NP/NOS $\leq$ 0.092 then ST = TRUE	If NP/AFA $\leq$ -0.433 then ST = TRUE	If NP/ASE $\leq$ 0.010 then ST = TRUE	If NP/AFA $\leq$ 0.017 then ST = TRUE
If FA/TA $\leq$ 0.504 and NP/NOS $\leq$ 0.172 and NICCE/NOS $\leq$ 0.097 then ST = TRUE	If NP/ATA $\leq$ 0.025 and NP(t)/NP(t - 1) $\leq$ 0.798 then ST = TRUE	If NOCF/CL $\leq$ 0.207 and SR/ACA > 0.350 and NP/NOS $\leq$ 0.092 then ST = TRUE	If NP(t)/NP(t - 1) $\leq$ -0.446 then ST = TRUE	If NP/NOS $\leq$ 0.020 then ST = TRUE	If NP/ASE $\leq$ 0.010 then ST = TRUE

**Table 8**  
Robust rules generated by DRSA.

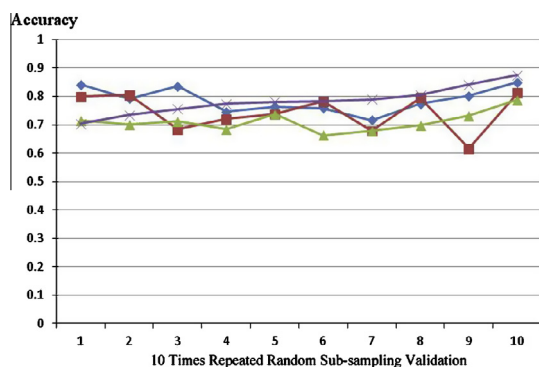
$t = 3$		$t = 4$		$t = 5$	
Ranked by coverage	Coverage	Ranked by coverage	Coverage	Ranked by coverage	Coverage
If $NP(t)/NP(t-1) \leq -0.369$ and $(CA-I)/CL1 \leq 2.125$ then $ST = TRUE$	29	If $MBC/AI \leq 14.349$ and $NP(t)/NP(t-1) \leq -0.310$ then $ST = TRUE$	34	If $FA/TA \leq 0.472$ and $NP/NOS \leq 0.022$ then $ST = TRUE$	25
If $NP/ATA \leq 0.018$ and $(CA-I)/CL1 \leq 1.141$ then $ST = TRUE$	26	If $NP/SR \leq 0.017$ and $NA/NOS \leq 3.267$ then $ST = TRUE$	32	If $NP/AFA \leq 0.015$ and $FA/TA \leq 0.559$ then $ST = TRUE$	22
If $(CA-I)/CL \leq 0.977$ and $EBIT/ATA \leq 0.032$ then $ST = TRUE$	24	If $CA/CL \leq 1.671$ and $NP/NOS \leq 0.040$ then $ST = TRUE$	27	If $NP/AFA \leq 0.016$ and $FA/TA \leq 0.472$ then $ST = TRUE$	21
If $NP/SR \leq 0.0118$ then $ST = TRUE$	22	If $NOCF/CL \leq 0.143$ and $NP/SR \leq 0.017$ then $ST = TRUE$	25	If $MBI(t)/MBI(t-1) \leq 0.186$ and $(CA-I)/CLO \leq -0.121$ then $ST = TRUE$	20
If $CA/TA \leq 0.755$ and $NP/NOS \leq 0.029$ then $ST = TRUE$	22	If $NP/SR \leq 0.010$ and $SR/AFA \leq 4.218$ then $ST = TRUE$	22	If $NP/SR \leq 0.008$ and $EBIT/ATA \leq 0.038$ then $ST = TRUE$	19

of ST with the condition that confidence value must be greater than 0.8 and the number of supporting instances must be greater than a quarter of the sample size. The top five rules ranked by confidence and coverage are presented in Table 7. Among the robust rules generated by DT, the criterion of  $NP/ATA \leq 0.025$  occurred in nine out of ten repeated trials. This indicated that the risk of financial distress was high when net profit of the firm was less than 2.5% of the value of average total assets held by the firm. Furthermore, DRSA was able to induce generalized decision rules as well. The results are shown in Table 8. Since the consistency (or confidence) level of DRSA was fixed at 1.0, the coverage of each rule was the only criterion that was used for choosing the rules. However, the coverage levels of DRSA rules were relatively low compared with DT rules. Among the top five robust rules generated by DRSA, the indicator NP/SR repeatedly occurred with an upper bound of 0.01. This indicated that the value of NP/SR should be closely monitored and when it dropped below 0.01 this could signal the possibility of ST warning.

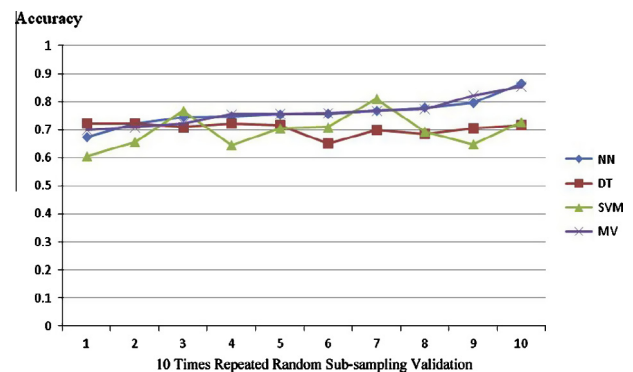
Multiple classifiers based on MV did not necessarily outperform the best single classifiers. With the exception of the cases with training to testing ratio of 80:20 in the 3-year time window and 60:40 in the 4-year time window, the average prediction accuracy of MV was always slightly lower than that of NN. Besides, the standard deviation of MV was larger than single classifiers as well. In other words, the multiple classifiers that are combined using MV did not improve the performance of NN significantly. This conclusion contradicted the result of Sun and Li (2008a), and Xiao et al. (2012). Multiple classifiers could overcome the weaknesses of single classifiers when single classifier candidates had similar competence in prediction of financial distress, as shown for 3-year and 4-year time window forecast with training to testing ratio at 60:40 (Figs. 4 and 5 respectively). However, when there was one model performing in a stable fashion at the best level like the case of the prediction for the 5-year time window, as shown in Fig. 6,

multiple classifiers that combined the best model with other models with poor performance lowered the predictive accuracy.

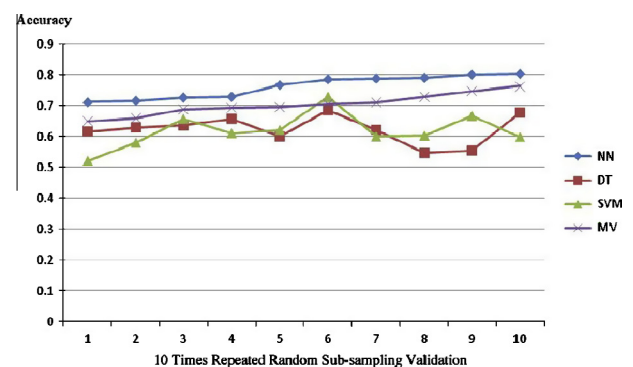
It seemed obvious that the closer we were to the year when the companies are labeled as ST, the higher was the accuracy of prediction. However, it was always better to detect the deterioration of listed companies' financial conditions as early as possible so that the managers got enough time and opportunity to take measures to avoid such a financial crisis. As the tiny signals of deterioration accumulated and magnified, the ability to predict financial distress became stronger but obviously managers got less time to react to the crisis. A tradeoff between prediction accuracy and time period of forecast was inevitable. The results of one-way ANOVA with Fisher's LSD procedure test indicated that forecasts in the three time periods were similar. Particularly, the prediction that is obtained in the 5-year time window was nearly as good as that



**Fig. 4.** Results using 10 times repeated random sub-sampling validation for classifiers in the 3-year time window.



**Fig. 5.** Results using 10 times repeated random sub-sampling validation for classifiers in the 4-year time window.



**Fig. 6.** Results using 10 times repeated random sub-sampling validation for classifiers in the 5-year time window.

of the 4-year time window when NN is used as the classifier (Fig. 4).

To answer the third question related to ascertaining the relative importance of the financial indicators, the input variables are selected according to a ranking method. In extant research, Altman (1968) selected five financial indicators of strong predictive ability from the initial set of 22 financial indicators using stepwise discriminant analysis. In this research, among the 31 indicators belonging to 6 categories, financial indicators reflecting capital expansion capacity, profitability, and business development capacity are found to be quite important in prediction of financial distress. We observed that in contrast to the significant financial indicators obtained by Altman (1984), debt and capital allocation turned out to be weak indicators because the target variable was set as ST rather than bankruptcy. The most important financial indicators that we obtained are related to net profit, earnings before income tax, cash flow, and net assets. This implied that profit and cash flow were important determinants of financial deterioration.

An interesting observation related to the 10 most important financial indicators in the different time periods of 3–5 years was that they were basically the same. For the 5-year time period the result of ranking for the top 7 inputs by scoring and by occurrence had very little divergence, whereas the same for the 3-year and the 4-year time periods turned out to be the same indicators but only different in the order of importance. Paired *t* test on the prediction accuracy before and after feature selection showed no significant difference for NN. This was an important finding as it implied that we could make equally effective prediction based on the top ten financial indicators instead of considering all 31 financial indicators.

## 6. Conclusion

In this study, data mining techniques are used to build models for predicting financial distress of Chinese listed companies based on 31 financial indicators. 107 ST companies and the same number of financially healthy listed companies are selected as the data set. The best method is obtained by comparing the performance of different prediction models at different time periods. The key contributions of this research are as detailed below.

We considered all companies that had been designated as ST during 2008–2011 to obtain a large sample of ST companies. Control samples are not confined to companies of the same size and belonging to the same industry, in order to make the data set less biased. The prediction ability of the financial indicators for the 3-year, 4-year, and 5-year time windows are compared to examine how early we could foresee the signs of financial deterioration at an acceptable accuracy. Financial data in the 5-year time window achieved a satisfactory predictive performance and even better than that in the 3-year time window.

Next, we adopted 31 financial indicators based on the Chinese accounting standards reflecting the solvency, profitability, operational capabilities, business development capacity, structural soundness, and capital expansion capabilities of the companies. In the feature selection procedure, one-way ANOVA *F* test is conducted to test the strength of relationship between each input variable and the target variable. Besides, we also evaluated the predictive ability of all financial indicators according to a ranking scheme using importance scores and frequency of occurrence. We obtained the top ten financial indicators with the highest predictive ability and they included return on total assets, net profit on total assets, earnings per share, and cash flow per share, among others. These indicators are found to have equivalent predictive abilities as the 31 indicators that are chosen initially

according to a paired *t* test on the predictive accuracy before and after feature selection.

Lastly, we identified NN as the best method for predicting financial distress in comparison to MV, DT, and SVM. The results of the 10 times repeated random sub-sampling validation and the impact of sensitivity analysis using multiple training to testing ratios strengthened the robustness of the findings. In terms of performance, NN based on the prune algorithm is found to be superior to that of the BPNN and the SVM, contrary to the finding reported by Ding et al. (2008).

Research on financial distress is usually limited to the use of financial indicators disclosed in accounting statements, thereby ignoring the role of non-financial indicators. In fact, some scholars have found that non-financial indicators such as corporate governance could be important in determining financial distress (Geng et al., 2011). In a future study, we would use non-financial indicators in the prediction model to improve the prediction accuracy. The existence of difference in accounting standards across countries might limit the generalizability of the results of this study for firms belonging to other countries. Future research could focus on building models for prediction of financial distress for different countries and compare and contrast between these models. The time window used to collect financial data of companies might affect the prediction accuracy significantly. We have criticized the ST risk prediction models based on financial data for 1-year and 2-year time windows, and compared the model performance of three different time windows in this paper. Prediction using data from 6-year to 10-year time windows could be taken up to decide how far back in time we could go to determine the signs of financial deterioration.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ejor.2014.08.016>.

## References

- Ahn, B. S., Cho, S. S., & Kim, C. Y. (2000). The integrated methodology of rough set theory and artificial neural network for business failure prediction. *Expert Systems with Applications*, 18(2), 65–74.
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589–609.
- Altman, E. I. (1984). The success of business failure prediction models: An international survey. *Journal of Banking & Finance*, 8(2), 171–198.
- Altman, E. I. (1993). *Corporate Financial Distress and Bankruptcy* (2nd ed.). New York: John Wiley & Sons.
- Altman, E. I., Heine, M. L., Zhang, L., & Yen, J. (2007). *Corporate financial distress diagnosis in China*. Salomon Center Working Paper, New York University.
- Bailey, W. B., Huang, W., & Yang, Z. (2011). Bank loans with Chinese characteristics: Some evidence on inside debt in a state-controlled banking system. *Journal of Financial and Quantitative Analysis*, 46(6), 1795–1830.
- Ball, R., & Foster, G. (1982). Corporate financial reporting: A methodological review of empirical research. *Journal of Accounting Research*, 20, 161–234.
- Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 4, 71–111.
- Błaszczyszński, J., Greco, S., & Słowiński, R. (2007). Multi-criteria classification – A new scheme for application of dominance-based decision rules. *European Journal of Operational Research*, 181(3), 1030–1044.
- Bose, I., & Pal, R. (2006). Predicting the survival or failure of click-and-mortar corporations: A knowledge discovery approach. *European Journal of Operational Research*, 174(2), 959–982.
- Carlos, S.-C. (1996). Self-organizing neural networks for financial diagnosis. *Decision Support Systems*, 17(3), 227–238.
- Chen, M. Y. (2013). A hybrid ANFIS model for business failure prediction utilizing particle swarm optimization and subtractive clustering. *Information Sciences*, 220, 180–195.
- Chen, W. S., & Du, Y. K. (2009). Using neural networks and data mining techniques for the financial distress prediction model. *Expert Systems with Applications*, 36(2), 4075–4086.
- Cohen, S., Doumpos, M., Neofytou, E., & Zopounidis, C. (2012). Assessing financial distress where bankruptcy is not an option: An alternative approach for local municipalities. *European Journal of Operational Research*, 218(1), 270–279.

- Dairui, L., & Jia, L. (2009). *Determinants of financial distress of ST and PT companies: A panel analysis of Chinese listed companies*. <[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1341795](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1341795)> Last accessed 07.03.14.
- Davis, J., & Goadrich, M. (2006). The relationship between precision-recall and ROC curves. In *Proceedings of the 23rd international conference on machine learning, Pittsburgh, USA* (pp. 233–240).
- Dhar, S., Mukherjee, T., & Ghoshal, A. K. (2010). Performance evaluation of neural network approach in financial prediction: Evidence from India Market. In *Proceedings of the international conference on communication and computational intelligence, Erode, India* (pp. 597–602).
- Dimitras, A. I., Zanakis, S. H., & Zopounidis, C. (1996). A survey of business failures with an emphasis on prediction methods and industrial applications. *European Journal of Operational Research*, 90(3), 487–513.
- Ding, Y., Song, X., & Zen, Y. (2008). Forecasting financial condition of Chinese listed companies based on support vector machine. *Expert Systems with Applications*, 34(4), 3081–3089.
- du Jardin, P., & Séverin, E. (2012). Forecasting financial failure using a Kohonen map: A comparative study to improve model stability over time. *European Journal of Operational Research*, 221(2), 378–396.
- Fan, A., & Palaniswami, M. (2000). Selecting bankruptcy predictors using a support vector machine approach. In *Proceedings of the international joint conference on neural networks, Como, Italy*.
- Fletcher, D., & Goss, E. (1993). Forecasting with neural networks: An application using bankruptcy data. *Information & Management*, 24(3), 159–167.
- Frydman, H., Altman, E. I., & Kao, D. L. (1985). Introducing recursive partitioning for financial classification: The case of financial distress. *Journal of Finance*, 40(1), 269–291.
- Geng, Z., Tan, L., Gao, X., Ma, Y., Feng, L., & Zhu, J. (2011). Financial distress prediction models of listed companies by using non-financial determinants in Bayesian criterion. In *Proceedings of the international conference on management and service science, Wuhan, China*.
- Gestel, T. V., Baesens, B., Suykens, J. A. K., Van den Poel, D., Baestaens, D.-E., & Willekens, M. (2006). Bayesian kernel based classification for financial distress detection. *European Journal of Operational Research*, 172(3), 979–1003.
- Greco, S., Matarazzo, B., Slowinski, R., & Stefanowski, J. (2001). Variable consistency model of dominance-based rough sets approach. In *Rough sets and current trends in computing* (pp. 170–181). Berlin Heidelberg: Springer.
- Greco, S., Matarazzo, B., Slowinski, R., & Zanakis, S. (2011). Global investing risk: A case study of knowledge assessment via rough sets. *Annals of Operations Research*, 185(1), 105–138.
- Jo, H., Han, I., & Lee, H. (1997). Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis. *Expert Systems with Applications*, 13(2), 97–108.
- Lau, A. (1987). A five-state financial distress prediction model. *Journal of Accounting Research*, 25(1), 127–138.
- Li, Z., Crook, J., & Andreeva, G. (2013). Chinese companies distress prediction: An application of data envelopment analysis. *Journal of the Operational Research Society*, 65(3), 466–479.
- Li, H., & Sun, J. (2009). Gaussian case based reasoning for business-failure prediction with empirical data in China. *Information Sciences*, 179(1–2), 89–108.
- Lin, F., Liang, D., Yeh, C. C., & Huang, J. C. (2014). Novel feature selection methods to financial distress prediction. *Expert Systems with Applications*, 41(5), 2472–2483.
- Martin, D. (1977). Early warnings of bank failure: A logit regression approach. *Journal of Banking and Finance*, 1(3), 249–276.
- McKee, T. E. (2000). Developing a bankruptcy prediction model via rough sets theory. *Intelligent Systems in Accounting, Finance & Management*, 9(3), 159–173.
- Meyer, P. A., & Pifer, H. W. (1970). Prediction of bank failures. *The Journal of Finance*, 25(4), 853–868.
- Ministry of Corporate Affairs (2014). Government of India. <[http://www.mca.gov.in/MCA21/default\\_comp\\_list.html](http://www.mca.gov.in/MCA21/default_comp_list.html)> Accessed 06.03.14.
- Newton, G. W. (1975). *Bankruptcy and insolvency accounting*. New York: The Ronald Press.
- Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109–131.
- Orr, K. G. (2003). *Corporate insolvency*. Times of Malta. <<http://www.timesofmalta.com/articles/view/20130422/local/Corporate-insolvency.466585>> Accessed 06.03.14.
- Palmeri, C. (2008). JPMorgan chase to buy Washington mutual. <<http://www.businessweek.com/stories/2008-09-26/jpmorgan-chase-to-buy-washington-mutualbusinessweek-business-news-stock-market-and-financial-advice>> Accessed 06.03.14.
- Ravi Kumar, P., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques – A review. *European Journal of Operational Research*, 180(1), 1–28.
- Ross, S., Westerfield, R., & Jaffe, J. (2000). *Corporate finance*. New York: McGraw-Hill Irwin.
- Sevim, C., Oztekin, A., Bali, O., Gumus, S., & Guresen, E. (2014). Developing an early warning system to predict currency crises. *European Journal of Operational Research*, 237(3), 1095–1104.
- Shearer, C. (2000). The CRISP-DM model: The new blueprint for data mining. *Journal of Data Warehousing*, 5(4), 13–22.
- Shie, F. S., Chen, M. Y., & Liu, Y.-S. (2012). Prediction of corporate financial distress: An application of the American banking industry. *Neural Computing and Applications*, 21(7), 1687–1696.
- Shin, K. S., & Lee, Y. J. (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications*, 23(3), 321–328.
- Sun, J., & Li, H. (2008a). Data mining method for listed companies' financial distress prediction. *Knowledge-Based Systems*, 21(1), 1–5.
- Sun, J., & Li, H. (2008b). Listed companies' financial distress prediction based on weighted majority voting combination of multiple classifiers. *Expert Systems with Applications*, 35(3), 818–827.
- Tax, D. M., Van Breukelen, M., Duin, R. P., & Kittler, J. (2000). Combining multiple classifiers by averaging or by multiplying? *Pattern Recognition*, 33(9), 1475–1485.
- Theodossiou, P., Kahya, E., Saidi, R., & Philippatos, G. (1996). Financial distress and corporate acquisitions: Further empirical evidence. *Journal of Business Finance & Accounting*, 23(5–6), 699–719.
- Tsai, C.-F. (2014). Combining cluster analysis with classifier ensembles to predict financial distress. *Information Fusion*, 16, 46–58.
- Usai, M. G., Goddard, M. E., & Hayes, B. (2009). LASSO with cross-validation for genomic selection. *Genetical Research*, 91(6), 427–436.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York: Springer-Verlag.
- Wanke, P., Barros, C. P., & Faria, J. R. (2014). Financial distress drivers in Brazilian banks: A dynamic slacks approach. *European Journal of Operational Research*. <http://dx.doi.org/10.1016/j.ejor.2014.06.044>.
- Wilson, R. L., & Sharda, R. (1994). Bankruptcy prediction using neural networks. *Decision Support Systems*, 11(5), 545–557.
- Xiao, Z., Yang, X., Pang, Y., & Dang, X. (2012). The prediction for listed companies' financial distress by using multiple prediction methods with rough set and Dempster-Shafer evidence theory. *Knowledge-Based Systems*, 26, 196–206.
- Xu, L., Krzyzak, A., & Suen, C. Y. (1992). Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE Transactions on Systems, Man and Cybernetics*, 22(3), 418–435.
- Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22, 59–82.
- Zopounidis, C., & Dimitras, A. I. (1998). *Multicriteria decision aid methods for the prediction of business failure*. New York: Springer.