Expert Systems with Applications 37 (2010) 2371–2376

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Beyond business failure prediction

Wei-Wen Wu *

Department of International Trade, Ta Hwa Institute of Technology, 1. Ta Hwa Road, Chiung-Lin, Hsin-Chu 307, Taiwan

ARTICLE INFO

Keywords:
Business failure
Prediction model
Automatic clustering
Feature selection

ABSTRACT

Business failures are depressing events which not only decimate the benefits of stakeholders but also affect the continuing development of economy and society. In order to reduce the impact of business failure, various models of business failure prediction have been developed. Although failure prediction models currently achieve a collective average accuracy of more than 85%, few persons can bear a risk of less than 100% accuracy under the present conditions of economic crisis. It is of particular interest that current failure prediction models have tended to adopt the technique of matching up failed and non-failed firms. This method, however, seems to have merely led to further complications. This paper proposes a method which directly explores the features of failed firms rather than researching pairs of failed and non-failed firms. To this end, automatic clustering techniques and feature selection techniques are employed for this study.

1. Introduction

After the demise of giant organizations like WorldCom and Enron, investors in global economies have become cautious of risks, especially risk of business failure (Aziz & Dar, 2006). Business failures are destructive events which not only harm the continuing development of economy and society, but also wipe out the benefits of a wide range of stakeholders, from stockholders, creditors, labor unions, governmental bodies, and employees, to customers and suppliers. The term "business failure" refers to a firm discontinuing its business operations because of its inability to make enough profit (Ahn, Cho, & Kim, 2000). A business failure may happen as a result of poor management skills, insufficient marketing, and lack of ability to compete with other similar businesses. It can also be the result of a domino effect caused by business failures of suppliers or customers. Depending on our specific concern or standpoint, business failure can be defined in a variety of ways. Under a broad definition, a business failure is a situation in which a firm cannot pay lenders, preferred stock shareholders, suppliers, etc., or a bill is overdrawn, or the firm is bankrupt according to law (Ahn et al., 2000; Dimitras, Zanakis, & Zopounidis, 1996). Referring to Shuai and Li (2005), corporate bankruptcy always brings huge economic losses to investors and others, together with a substantial social and economical cost to a nation. This paper focuses on two types of business failure, all of them exemplified in cases of listed companies which went into bankruptcy or delisting.

In order to prevent business failure and the ensuing damage, various methods of business failure prediction have been presented. Kamath and He (2006) state that since the pioneering work by Beaver (1967) and Altman (1968), numerous contributions have been made to the development and refinement of failure prediction models. Aziz and Dar (2006) divide failure prediction models into three categories: classical statistical models, artificially intelligent expert system models, and theoretical models. These failure prediction models are important and meaningful for the following purposes: (1) corporate managers can employ failure prediction methods to develop early warning systems on imminent business failure, and may then take appropriate actions to prevent such failures; (2) sponsors and financial institutions can utilize failure prediction models to enable better decision-making in evaluating firms to collaborate with; (3) investors can apply failure prediction models for selecting healthy firms to invest in; and (4) prospective employees can make use of failure prediction methods when screening for robust firms to work for.

Referring to Kamath and He (2006), currently failure prediction models reach a collective average accuracy of more than 85%. However, few persons can tolerate an accuracy of less than 100% under conditions brought about by the greatest economic crisis in almost 80 years. As a result of the US subprime crisis in 2007, and the ruin of global banking giants like Lehman Brothers, many national economies are hurtling toward recession. How to recognize imminent business failure, and thus beat the economic crunch and achieve survival, has become an imperative issue. Because business failure seriously affects investment performance and job opportunities, we need to seek a more effective way to identify and handle it. To this end, a different viewpoint or research approach is called for.

Generally, the current approach to business failure is inclined to follow the traditional method of pairing up failed and non-failed firms. In other words, the existing approach to failure prediction

* Tel.: +886 3 5927700x2902; fax: +886 3 5925715.
E-mail address: itmike@thit.edu.tw

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.07.056

focuses on distinguishing failed firms from non-failed firms. However, due to the difficulties involved in the prerequisite of matching up two diverse groups, it is hard to approach a classification accuracy of 100% using current methodologies. More importantly, it is clear that the crucial factor, when we are attempting to limit the damage caused by business failures, is to identify the features of failed firms rather than to argue the extent of classification accuracy. In this situation, it is better to directly explore the features of failed firms rather than to investigate numerous pairs of failed and non-failed firms.

An additional important factor is financial ratios. These are mainly taken from a firm's balance sheet, income statement, cash flow statement, and so on. Financial ratios can be used not only to evaluate the overall financial condition and profitability of a firm, but also to compare the strengths and weaknesses between various firms. Very often, financial ratios are also applied to build failure prediction models. Hence, this paper proposes a method to explore the features of failed firms' financial ratios using automatic clustering techniques and feature selection techniques. The remainder of this paper is organized as follows. In Section 2, the business failure and prediction methods are discussed. In Section 3, research framework and methods are presented. In Section 4, an empirical study is conducted to illustrate implementation of the proposed method. Finally, from the findings of this research, we can derive some conclusions and implications for management.

2. Business failure and prediction methods

2.1. Business failure

The health of firm in a highly competitive business environment is dependent upon its capability of achieving profitability and financial solvency. This means that a firm becomes unhealthy, or deteriorates to the point where it is in danger of suffering business failure, when it loses its competence to maintain profitability and financial solvency. Business failure is not only common with new start-ups but also with listed companies, and it can easily happen to firms of any and all sizes. According to Ooghe and Waeyaert (2004), the causes of bankruptcy can be grouped into five interactive aspects. These include general environment (economics, technology, foreign countries, politics, and social factors), immediate environment (customers, suppliers, competitors, banks and credit institutions, stockholders, and misadventure), management/entrepreneur characteristics (motivation, qualities, skills, and personal characteristics), corporate policy (strategy and investments, commercial, operational, personnel, finance and administration, corporate governance), and company characteristics (size, maturity, industry, and flexibility). Furthermore, Ooghe and De Prijcker (2008) note that there are four types of failure process: the failure process of unsuccessful startups, the failure process of ambitious growth companies, the failure process of "dazzling growth" companies, and the failure process of apathetic established companies. In all these cases of business failure, eventually the firms involved cannot help but show their poor business performances in financial statements from which useful financial ratios can be calculated to help us evaluate the extent of failure.

2.2. Failure prediction methods

Failure prediction models have traditionally been based on estimates using data derived from the matching up of failed and non-failed firms from different industries. The validity of the resulting models is then evaluated by classifying firms into failed class and non-failed class (Kamath & He, 2006). Several previous studies have made contributions by providing a detailed review of these models in the domain of failure prediction. Shuai and Li (2005) mention that (1) following the work of Altman (1968) that employs multivariate discriminant analysis to differentiate between failed and non-failed firms, several methods, such as logit analysis, probit analysis, and linear programming, have been proposed; (2) however, those conventional statistical methods have practical limitations owing to their restrictive assumptions such as the linearity, normality and independence among predictor or input variables; and (3) to overcome the limitations of such statistical assumptions, a number of new techniques have been developed. These include "expert systems", neural networks, rough set theory, and genetic programming. According to Huang, Tsai, Yen, and Cheng (2008), (1) earlier studies related to failure prediction mostly use statistical methods such as multiple discriminant analysis, regression analysis, and linear discriminant analysis; (2) recently, it has become popular to apply artificial intelligence and machine learning techniques to the issue of failure prediction; and (3) several studies show that machine learning models outperform traditional statistical models.

Kumar and Ravi (2007) divide failure prediction techniques into two broad categories: statistical and intelligent techniques. Statistical techniques are such techniques as linear discriminant analysis, multivariate discriminate analysis, quadratic discriminant analysis, logistic regression and factor analysis. Intelligent techniques include neural network architectures (multi-layer perception, probabilistic neural networks, auto-associative neural network, self-organizing map, learning vector quantization, and cascade correlation neural network), decision trees, case-based reasoning, evolutionary approaches, rough sets, soft computing (hybrid intelligent systems), operational research techniques (linear programming, data envelopment analysis, quadratic programming), support vector machine, and fuzzy logic techniques. Aziz and Dar (2006) conduct an extensive literature review and report that (1) failure prediction models can be separated into three categories: statistical models, artificially intelligent expert system models (AIES), and theoretical models; and (2) the average overall predictive accuracy (one year before actual bankruptcy) achieved by these models is somewhat over 85%.

3. Research framework and methods

As indicated by the above discussion current studies related to business failure prediction are primarily concerned with the limitations of statistical assumptions, and the classification accuracies of prediction models. Dimitras et al. (1996) note that (1) even a good prediction model cannot successfully predict business failures across different types of firms; and (2) the usefulness of a prediction model is limited and unequal in different countries, sectors and periods of time. They note that prediction methods usually comprise three parts: sampling and data collection; method selection and specification of variables to develop a predictive model; and model validation. Shuai and Li (2005) highlight the significance of non-financial information, in cases, for example, of abnormal changes such as changing the CEO, the financial manager or the auditor in the year previous to the failure. Additionally, Dimitras et al. (1996) emphasize the importance of non-financial factors for business failure prediction. These factors include management, personnel, products, equipment, and so on.

Academic researchers have developed various modeling techniques, each having distinct assumptions and specific computational complexities (Balcaen & Ooghe, 2006). Tsakonas, Dounias, Doumpos, and Zopounidis (2006) also note that failure prediction needs to be conducted not only correctly but in a timely fashion. Furthermore, Chava and Jarrow (2004) stress that most of the case studies in the existing literature employ only yearly observations, while failure prediction is markedly improved using monthly observation intervals. Grice and Ingram (2001) examine the issue of whether Altman's *Z*-score model is still helpful for evaluating a firm's financial health, and report that (1) Altman's original model is not effective in classifying firms in recent periods; (2) this is particularly true in the case of predicting the business failure of non-manufacturing firms. Additionally, Kamath and He (2006) indicate that data collected from several different industries to develop a model is likely to ignore the heterogeneity of the observations, and therefore risks introducing bias in the estimation of a model's parameters.

Although all these studies have raised questions about current prediction methods, they are all the same in that they are followers of the classical paradigm. The classical paradigm of failure prediction models has some established features. One of these is that they are required to use data which comes from a large number of pairs of failed and non-failed firms. Another is that the performance of a prediction model is focused on classification accuracy, for example Cielen, Peeters, and Vanhoof (2004), Lee, Booth, and Alam (2005), Shin, Lee, and Kim (2005), and Thomas (2007), among others. This gives rise to the question: why not directly shine the spotlight on the failed firms. The technique of matching up failed and non-failed firms actually makes problems more complicated. Additionally, it is apparent that grasping the features of failed firms' financial ratios is more important than arguing the classification accuracies among prediction models. Hence, this paper sets out to straightforwardly explore the features of failed firms' financial ratios.

Beaver (1967) and Altman (1968) initially applied financial ratios in building failure prediction models. Since then, the selection and evaluation of financial ratios has become a vital element of failure prediction methods. Financial ratios can be categorized according to several aspects in order to measure the business performance or competence of a firm. For example, financial ratios can be used to measure a firm's profitability, liquidity, capital structure, and efficiency. According to Huang et al. (2008), financial ratios are important tools in prediction of business failures, as well as are commonly used to develop the models or classifiers. As shown in Fig. 1, this paper targets failed firms and aims to seek out relevant features of their financial ratios. To this end, automatic clustering techniques are employed to automatically divide targeted failed firms into some clusters according to characteristics of financial ratios. In order to simplify the task of analysis, as well as to increase the classification accuracy, feature selection techniques are used to reduce the overall number of financial ratios analyzed. Also, this paper particularly emphasizes the importance of both expert knowledge and data mining techniques in feature selection. This means that it is preferable to conduct the analysis task using not only the data mining technique but also the expert knowledge, and to compare their performances of classification accuracies in terms of the feature selection. In this way, more accurate results and practical insights can be obtained.

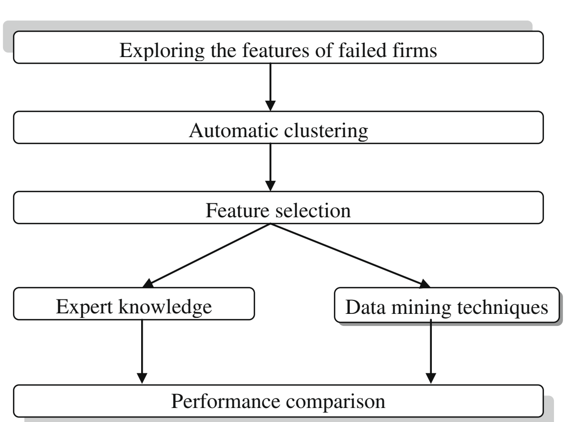

Fig. 1. Research framework.

4. Empirical study

4.1. Financial ratios and dataset

Business failure does not happen overnight and therefore a longitudinal and holistic perspective is needed (Ooghe & De Prijcker, 2008). Referring to Huang et al. (2008), financial ratios can be classified into five categories: 'Financial structure', 'Credit standing', 'Operating standing', 'Profitability', and 'Short-term credit standing'. For this study, quarterly financial ratios of failed firms were collected from the Financial Data Banks of Taiwan Economic Journal (TEJ). For purposes of thoroughly investigating the features of failed firms' financial ratios, this study acquired data on 175 failed firms (from various business sectors, but excluding the banking and finance industry) ranging from the fourth quarter of 1990 to the second quarter of 2008. For the sake of the data completeness, 163 failed firms were used in this study.

As shown in Table 1, there were a total of 15 financial ratios selected for use in this study: Return on Total Assets (ROA), Gross margin, Operating margin, Net profit margin, Cash flow Ratio (CF Ratio), Current Ratio, Quick Ratio, Debt Ratio, Shareholder's Equity/Total Assets Ratio (ET Ratio), Long-term Capital/Fixed Assets Ratio (LF Ratio), Interest Coverage Index (IC Index), Total Assets Turnover Ratio (TAT Ratio), Accounts Receivable Turnover Ratio (ART Ratio), Inventory Turnover Ratio (IT Ratio), and Fixed Assets Turnover Ratio (FAT Ratio).

4.2. Automatic clustering and feature selection

To implement the automatic clustering and the feature selection, the software known as WEKA was used here because it is free and easy to operate. WEKA is a package which incorporates a large number of machine learning algorithms for data mining tasks, and it provides a series of comprehensively practical utilities, such as: Preprocess, Classify, Cluster, Associate, Select attributes, Visualize.

For this study, three automatic clustering algorithms (EM, *X*-means, FilteredClusterer) were available. Referring to the WEKA manual, the Expectation Maximization (EM) clustering algorithm assigns a probability distribution to each instance which indicates the probability of it belonging to each of the clusters; the *X*-means clustering algorithm extends *K*-means with efficient estimation of the number of clusters; and the FilteredClusterer clustering algorithm is a meta-clusterer which offers the possibility to apply filters directly before the clusterer is learned. In order to compare performances of classification accuracies among three automatic clustering algorithms, eighteen classification algorithms were used here, including NaiveBayes, HillClimber, K2, SimulatedAnnealing, TabuSearch, TAN (Tree Augmented Naive Bayes), Multinomial Logistic, Multilayer Perception, RBFNetwork, Simple Logistic, SMO (Sequential minimal optimization), IB1 (Nearest-neighbour), IBk (*K*-nearest neighbours), LogitBoost, J48 (the WEKA version of C4.5), RandomForest, DTNB (a decision table/naive bayes hybrid classifier), and PART (a rule generator using J48 to generate pruned decision trees). Table 2 exhibits the result of comparative performances. With the aspect of geometric mean, the FilteredClusterer has the best classification accuracy, at 91.50%. This means that the sample of 163 failed firms is best divided into two classes.

According to result of the FilteredClusterer, the feature selection was conducted with two solutions: Solution 1 (based on data min-

Table 1
Financial ratios.

Categories	Financial ratios
Financial structure	Shareholder's equity/total assets ratio
	Debt ratio (total liability/total assets)
	Long-term capital/fixed assets ratio
Credit standing	Current ratio (current assets/current liabilities)
	Quick ratio (current assets – [inventories + prepayments])/current liabilities)
	Interest coverage index (earnings before interest and taxes/interest expenses)
Operating standing	Accounts receivable turnover ratio (credit sales/average receivables)
	Inventory turnover ratio (cost of goods sold/average inventory)
	Fixed assets turnover ratio (sales/average fixed assets)
	Total assets turnover ratio (sales/average total assets)
Profitability	Gross margin (operating earnings/sales)
	Operating margin (operating income/sales)
	Net profit margin (net profits after taxes/sales)
	Return on total assets (net profits after taxes/total assets)
Short-term credit standing	Cash flow ratio (operating cash flow/current liabilities)

Table 2
Comparative performance of automatic clustering algorithms.

	EM (4 classes) (%)	*X*-means (3 classes) (%)	FilteredClusterer (2 classes) (%)	Geometric mean (%)
NaiveBayes	95.09	66.26	74.85	77.84
HillClimber	88.96	78.53	85.89	84.34
K2	88.96	78.53	85.89	84.34
SimulatedAnnealing	86.50	77.91	91.41	85.09
TabuSearch	88.96	78.53	85.89	84.34
TAN (Tree Augmented Naive Bayes)	90.18	80.37	93.25	87.76
Multinomial logistic	89.57	90.80	97.55	92.57
Multilayer perception	87.73	93.25	99.39	93.34
RBFNetwork	93.25	77.91	91.41	87.25
Simple logistic	91.41	95.71	98.77	95.25
SMO (sequential minimal optimization)	76.69	85.28	88.96	83.48
IB1 (nearest-neighbour)	84.05	88.34	95.09	89.05
IBk (*K*-nearest neighbours)	84.05	88.34	95.09	89.05
LogitBoost	91.41	85.89	94.48	90.52
J48	88.96	83.44	92.64	88.27
RandomForest	90.18	88.96	93.87	90.98
DTNB	92.64	80.37	92.03	88.16
PART	93.25	78.53	93.87	88.25
Geometric mean	88.89	82.87	91.50	87.68

Table 3
The result of feature selection.

CfsSubsetEval: ROA, Operating margin, Net profit margin, Debt Ratio
ConsistencySubsetEval: ROA, Gross margin, Operating margin, Net profit margin, Debt Ratio, TAT Ratio
ChiSquaredAttributeEval: Net profit margin (123.70), ROA (63.90), Operating margin (61.40), Gross margin (42.75), IC Index (40.28), Debt Ratio (30.77), ET Ratio (30.77), TAT Ratio (28.99), ART Ratio (21.89), Quick Ratio (21.00)
Filtered AttributeEval: Net profit margin (0.59), ROA (0.27), Operating margin (0.26), IC Index (0.19), Gross margin (0.17), ET Ratio (0.14), Debt Ratio (0.14), TAT Ratio (0.13), Quick Ratio (0.11), Current Ratio (0.11), ART Ratio (0.09)
GainRatio AttributeEval: Net profit margin (0.52), ROA (0.31), Operating margin (0.30), Gross margin (0.25), IC Index (0.19), ART Ratio (0.15), Debt Ratio (0.14), ET Ratio (0.14), TAT Ratio (0.13), Current Ratio (0.12), Quick Ratio (0.12)
InfoGainAttributeEval: Net profit margin (0.59), ROA (0.27), Operating margin (0.26), IC Index (0.19), Gross margin (0.17), ET Ratio (0.14), Debt Ratio (0.14), TAT Ratio (0.13), Quick Ratio (0.11), Current Ratio (0.11), ART Ratio (0.09)
ReliefF AttributeEval: ROA (0.12), TAT Ratio (0.06), Operating margin (0.03), Gross margin (0.02), Debt Ratio (0.02), ET Ratio (0.02), Net profit margin (0.04), ART Ratio (0.01), CF Ratio (0.01), Current Ratio (0.01)
Sy mmetricalUncertAttributeEval: Net profit margin (0.60), ROA (0.32), Operating margin (0.31), Gross margin (0.22), IC Index (0.21), Debt Ratio (0.15), ET Ratio (0.15), TAT Ratio (0.14), Quick Ratio (0.13), Current Ratio (0.12), ART Ratio (0.12)

ing techniques), and Solution 2 (based on expert knowledge). For Solution 1, eight data mining techniques were employed for the feature selection, including CfsSubsetEval (evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature), ConsistencySubsetEval (evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature), EvalChiSquaredAttributeEval (evaluates the worth of an attribute by computing the value of the chi-squared statistic with respect to the class), FilteredAttributeEval (class for running an arbitrary attribute evaluator on data that has been passed through an arbitrary filter), GainRatioAttributeEval (evaluates the worth of an attribute by measuring the gain ratio with respect to the class), InfoGainAttributeEval (evaluates the worth of an attribute by measuring the information gain with respect to the class), ReliefFAttributeEval (evaluates the worth of an attribute by repeatedly sampling an instance and considering the value of the given attribute for the nearest instance of the same and different class), and SymmetricalUncertAttributeEval (evaluates the worth of an attribute by measuring the symmetrical uncertainty with respect to the class). As shown in Table 3, we can see that four financial ratios (ROA, Operating margin, Net profit margin, Debt Ratio) were more significant than other financial ratios across these eight data mining techniques. These four financial ratios were regarded as the result of the feature selection for Solution 1.

As for expert knowledge about financial ratios analysis, the DuPont model is the most representative and comprehensive because it integrates key elements of the Income Statement with those of the Balance Sheet. The DuPont formula can be expressed as ROA = Net profit margin x Total Assets Turnover Ratio (TAT Ratio). It should be noted that Debt Ratio is very often used by experts to measure a firm's financial structure. Hence, these three financial ratios (Net profit margin, TAT Ratio, Debt Ratio) were viewed as the result of the feature selection for Solution 2. Further, the sample of 163 failed firms with three financial ratios (Net profit margin, TAT Ratio, Debt Ratio) was, respectively clustered by three automatic clustering algorithms (EM, *X*-means, FilteredClusterer). Correspondingly, Solution 2a (expert knowledge + EM), Solution 2b (expert knowledge + *X*-means), and Solution 2c (expert knowledge + FilteredClusterer) were obtained.

Table 4 exhibits the comparative performance of different solutions with eighteen classification algorithms. From the standpoint of geometric mean, Solution 1, with 91.59%, was slightly better than the original Solution, with 91.50%, that was reached without using feature reduction. Moreover, in addition to Solution 2a (88.24%), both Solution 2b and Solution 2c (88.24%) were better than Solution 1. Overall, Solution 2 outperformed Solution 1. This means that an expert knowledge-based approach surpasses a data mining technique-based approach in feature selection.

Of especial interest is that the result of our analysis reveals that the initial 15 financial ratios can be reduced into three financial ratios (Net profit margin, TAT Ratio, Debt Ratio) for purposes of identifying the features related to business failure. For example with regard to Solution 2c (see Table 5), either class1 ($n = 43$) or class2 ($n = 120$) showed extremely poor performances in each of three financial ratios. This is also true with regard to the average financial ratios for one quarter prior to business failure (see Table 6). Thus, the findings reveal that when we are seeking to determine the crucial features of business failure, we have no need for difficult prediction techniques, but rather that these three financial ratios deriving from expert knowledge are needed.

Table 7, referring to the Data Banks of Taiwan Institute of Economic Research (TIER), shows a series of average financial ratios

Table 5
Average financial ratios for business failure.

Class	Net profit margin (%)	Debt ratio (%)	TAT ratio
1	−22.76	71.72	0.28
2	−282.74	68.46	0.08

Table 6
Average financial ratios for one quarter prior to business failure.

Class	Net profit margin (%)	Debt ratio (%)	TAT ratio
1	−12.30	69.70	0.28
2	−169.52	63.52	0.09

from the manufacturing industry which can be used as guidelines to measure business failure. For example, from the perspective of the geometric mean, we may consider that a firm is walking the road of business failure if its Net profit margin, TAT Ratio, and Debt Ratio are lower than 3.55%, 0.94, and 49.36%, respectively. More importantly, if a firm has the features shown in Table 5 or Table 6, we need not hesitate, and can immediately judge that it is bound for business failure without employing any prediction techniques.

5. Implications and conclusions

The ability to identify or judge firms headed for failure is important for stakeholders at all times, and particularly in times of economic crisis. Current failure prediction models are based on the traditional approach, which proceeds by making pairs of failed and non-failed firms. Using this current approach, it is not possible to reach a classification accuracy of 100% due to the prerequisite of matching up two diverse groups. We believe that few persons can bear a risk of less than 100% accuracy under conditions of the greatest economic crisis in almost 80 years.

Furthermore, this paper questions the need of the method which proceeds by making pairs of failed and non-failed firms. In fact, we never need to distinguish a sour apple from a sweet apple. What we need to do is just to take a bite. Similarly, to identify firms at risk of failure, we just need to see the key financial ratios rather than to conduct failure predictions. When a person has caught a cold, the urgent thing is to get him/her to see a doctor rather than to make comparisons with a person has not caught a cold. In the

Table 4
Comparative performance of different solutions.

	Solution 2a (%)	Solution 2b (%)	Solution 2c (%)	Geometric mean (%)	Solution 1 (%)
NaiveBayes	93.87	84.05	84.05	87.20	91.41
HillClimber	89.57	100.00	100.00	96.39	92.03
K2	89.57	100.00	100.00	96.39	92.03
SimulatedAnnealing	88.96	100.00	100.00	96.18	93.25
TabuSearch	89.57	100.00	100.00	96.39	92.03
TAN (Tree Augmented Naive Bayes)	90.18	100.00	100.00	96.61	92.03
Multinomial logistic	88.34	99.39	99.39	95.56	92.64
Multilayer perception	84.05	99.39	99.39	93.99	90.80
RBFNetwork	85.89	98.16	98.16	93.89	89.57
Simple logistic	88.34	99.39	99.39	95.56	94.48
SMO (Sequential minimal optimization)	74.23	91.41	91.41	85.28	84.05
IB1 (Nearest-neighbour)	87.73	96.93	96.93	93.76	90.18
IBk (*K*-nearest neighbours)	87.73	96.93	96.93	93.76	90.18
LogitBoost	93.25	100.00	100.00	97.70	92.64
J48	89.57	100.00	100.00	96.39	93.25
RandomForest	91.41	100.00	100.00	97.05	93.25
DTNB	88.96	100.00	100.00	96.18	92.64
PART	88.96	100.00	100.00	96.18	92.64
Geometric mean	88.24	98.00	98.00	94.64	91.59

Table 7
Average financial ratios of manufacturing industry.

	1999	2000	2001	2002	2003	2004	2005	2006	Geometric mean
Current ratio (%)	120.65	121.49	121.19	124.89	131.66	132.11	137.98	165.41	131.25
Quick ratio (%)	74.01	75.02	75.83	83.44	89.00	86.68	93.33	121.27	86.27
Debt ratio (%)	51.18	50.88	50.55	50.46	48.90	49.18	47.61	46.36	49.36
ET ratio (%)	48.82	49.12	49.45	49.54	51.10	50.82	52.39	53.64	50.58
LF ratio (%)	154.79	155.88	158.92	164.67	178.10	177.47	191.89	203.59	172.38
TAT ratio (times)	0.84	0.86	0.74	0.83	1.55	0.89	0.94	1.05	0.94
ART ratio (times)	5.59	5.78	5.30	5.77	10.64	6.06	6.08	6.61	6.32
IT ratio (times)	4.81	5.03	4.69	5.51	11.17	5.99	6.18	7.30	6.09
FAT ratio (times)	2.20	2.21	1.88	2.16	4.34	2.51	2.76	3.24	2.57
ROA (%)	2.80	3.72	0.45	2.09	6.85	5.61	5.86	6.91	3.33
ROE (%)	5.78	7.60	0.91	4.23	13.40	11.01	11.36	13.03	6.62
Gross margin (%)	16.67	17.21	15.58	16.30	16.24	17.09	14.93	13.57	15.91
Operating margin (%)	4.52	5.28	2.89	4.36	5.25	7.26	5.78	6.12	5.02
Net profit margin (%)	3.34	4.34	0.61	2.53	4.41	6.29	6.22	6.56	3.55

same way, this paper suggests we proceed to directly discover the features of failed firms' financial ratios rather than researching pairs of failed and non-failed firms.

These are bad times, it is true, but it is also a good time to adopt a different viewpoint in research on the phenomenon of business failure. The author believes that from the results of the empirical study made according to the proposed method, useful and relevant findings are obtained. Firstly, if we wish to prevent the impact of business failure, it is appropriate to focus on failed firms themselves, not on pairs of failed and non-failed firms. Moreover, to grasp the features of failed firms' financial ratios is more meaningful than to argue the classification accuracies among prediction models. Indeed, identifying a failed firm is not so difficult. This is because we can identify a failed firm by looking at certain key financial ratios at a point in time one quarter prior to business failure. Furthermore, it is advantageous to reduce numerous financial ratios into a few key financial ratios in order to simply and effectively conduct the analysis task. To this end, feature selection is needed. As for feature selection, the result of a comparative performance analysis shows that an expert knowledge-based approach surpasses a data mining technique-based approach. This reveals that, while data mining techniques are powerful, expert knowledge is even more valuable and indispensable.

Business failure is an event that fatally affects our investment performance and reduces job opportunities. Recognizing and handling the business failure problem in a timely manner is becoming a required competency for survival in the economic crunch. This study succeeds in uncovering the essential features of failed firms' financial ratios, and constitutes a new and different approach, going beyond the traditional methodologies of business failure prediction. With regard to future research, international comparative studies are the next challenge which should be taken up.

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. *Journal of Finance, 23*(4), 589–609.

Aziz, M. A., & Dar, H. A. (2006). Predicting corporate bankruptcy: Where we stand? *Corporate Governance, 6*(1), 18–33.

Balcaen, S., & Ooghe, H. (2006). 35 years of studies on business failure: An overview of the classic statistical methodologies and their related problems. *British Accounting Review, 38*(1), 63–93.

Beaver, W. (1967). Financial ratios predictors of failure. Empirical research in accounting: Selected studies 1966. *Journal of Accounting Research, 4*(Suppl.), 71–111.

Chava, S., & Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. *Review of Finance, 8*(4), 537–569.

Cielen, A., Peeters, L., & Vanhoof, K. (2004). Bankruptcy prediction using a data envelopment analysis. *European Journal of Operational Research, 154*(2), 526–532.

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.

Grice, J. S., & Ingram, R. W. (2001). Tests of the generalizability of Altman's bankruptcy prediction model. *Journal of Business Research, 54*(1), 53–61.

Huang, S. H., Tsai, C. F., Yen, D. C., & Cheng, Y. L. (2008). A hybrid financial analysis model for business failure prediction. *Expert Systems with Applications, 35*(3), 1034–1040.

Kamath, R., & He, Y. (2006). Business failure prediction in the retail industry: An empirical evaluation of generic prediction models. *Academy of Accounting Studies Journal, 10*(2), 97–110.

Kumar, P. R., & 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.

Lee, K., Booth, D., & Alam, P. (2005). A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms. *Expert Systems with Applications, 29*(1), 1–16.

Ooghe, H., & De Prijcker, S. (2008). Failure processes and causes of company bankruptcy: A typology. *Management Decision, 46*(2), 223–242.

Ooghe, H., & Waeyaert, N. (2004). Oorzaken van faling: Literatuuroverzicht en conceptueel verklaringsmodel. *Economisch en Sociaal Tijdschrift, 57*(4), 367–393.

Shin, K. S., Lee, T. S., & Kim, H. J. (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications, 28*(1), 127–135.

Shuai, J. J., & Li, H. L. (2005). Using rough set and worst practice DEA in business failure prediction. *Lecture Notes in Computer Science, 3642*(2), 503–510.

Thomas, E. M. (2007). Altman's 1968 bankruptcy prediction model revisited via genetic programming: New wine from an old bottle or a better fermentation process? *Journal of Emerging Technologies in Accounting, 4*(1), 87–101.

Tsakonas, A., Dounias, G., Doumpos, M., & Zopounidis, C. (2006). Bankruptcy prediction with neural logic networks by means of grammar-guided genetic programming. *Expert Systems with Applications, 30*(3), 449–461.