Journal of
Marine Science and Technology

Volume 18 | Issue 4 Article 4

THE ROLE OF NON-FINANCIAL FEATURES RELATED TO CORPORATE GOVERNANCE IN BUSINESS CRISIS PREDICTION

Fengyi Lin
Department of Business Management, National Taipei University of Technology

Deron Liang
Department of Computer Science and Information Engineering, National Central University, Taoyuan County, Taiwan, R.O.C.

Wing-Sang Chu
Department of Computer Science and Information, National Taiwan Ocean University, Keelung, Taiwan, R.O.C

Recommended Citation
Lin, Fengyi; Liang, Deron; and Chu, Wing-Sang (2010) "THE ROLE OF NON-FINANCIAL FEATURES RELATED TO CORPORATE GOVERNANCE IN BUSINESS CRISIS PREDICTION," *Journal of Marine Science and Technology*: Vol. 18: Iss. 4, Article 4.
DOI: 10.51400/2709-6998.1901
Available at: https://jmstt.ntou.edu.tw/journal/vol18/iss4/4

THE ROLE OF NON-FINANCIAL FEATURES RELATED TO CORPORATE GOVERNANCE IN BUSINESS CRISIS PREDICTION

Fengyi Lin*, Deron Liang**, and Wing-Sang Chu***

Key words: corporate governance, financial prediction, support vector machines.

ABSTRACT

Recent outbreak of corporate financial crises worldwide has brought attention to the need for a new international financial architecture which rests on crisis prediction and crisis management. It is therefore both desirable and vital to explore new predictive techniques for providing early warnings aganist bankruptcy. Financial data have been widely used by researchers to predict financial distress or business crisis, but few studies exploit the use of non-financial indicators related to corporate governance to construct business crisis prediction model. This article introduces into the field of business crisis prediction model based on a combination of both financial and corporate governance related non-financial data. The experiment results show that the combined use of both financial and non-financial features with SVM model leads to a more accurate prediction of financial distress.

I. INTRODUCTION

Financial prediction is not only an important but also a challenging problem that generates extensive studies over the past decades. Recent outbreak of corporate financial crises worldwide has intensified the need to reform the existing financial architecture. It is generally believed that symptoms and alarms can be observed prior to a business encounters financial difficulty or crisis. The overall objective of business crisis prediction is to build models that can extract knowledge of risk evaluation from past observations and to apply it to evaluate business crisis risk of companies with a much broader scope. Eichengreen [13] identifies the policies of the new international financial architecture as crisis prevention, crisis prediction and crisis management. Financial indicators have been consulted by researchers as a major basis for predicting financial distress and business crises while other common methodologies include peer group analysis, comprehensive risk assessment systems, and statistical and econometric models [24].

Yeh and Woidtke [32] suggest that corporate governance factors, such as corporate board structure, concentrated ownership and shareholder concentration, should be taken into consideration when measuring the possibility of bankruptcy. Several recent financial scandals in Taiwan were characterized by the common trait of shareholding of board members, ratio of pledged shares of board members, and frequent changes in certified public accountants (CPAs) by distressed companies prior to bankruptcy. We have therefore included non-financial features related to corporate governance in our proposed classification model.

Recently, many researchers have endeavored to construct automatic classification systems by using data mining methods, such as statistical models and artificial intelligence (AI) techniques [3, 9, 12]. The former include linear regression, linear multivariate discriminant analysis (MDA), logit analysis and multidimensional scaling while the latter consist mainly of back propagation neural networks and case base reasoning. In addition to these classification methods, the support vector machine (SVM) proposed by Boster, Guyon, and Vapnik [3] has been successfully applied to many areas, including financial time series forecasting, credit scoring, and drug design [5]. While ANN implements empirical risk minimization principle to minimize the error on training set, SVM utilizes structural risk minimization principle to minimize generalization error. Therefore, the solution of SVM may be global optimum while ANN tends to fall into local optimum [12]. However, only few researchers have adopted SVM to examine non-financial features related to corporate governance for predicting corporate financial distress. Therefore, our study attempts to identify potential predictors to help users identify underlying characteristics of distressed firms.

The aim of this paper is twofold. First, this paper explores not only the role of financial features but also the role of

Paper submitted 05/02/08; revised 06/05/09; accepted 07/17/09. Author for correspondence: Deron Liang (e-mail: drliang@csie.ncu.edu.tw).
**Department of Business Management, National Taipei University of Technology*
***Department of Computer Science and Information Engineering, National Central University, Taoyuan County, Taiwan, R.O.C.*
****Department of Computer Science and Information, National Taiwan Ocean University, Keelung, Taiwan, R.O.C.*

non-financial features in business crisis prediction. For this purpose we examine empirically whether the combined consideration of both financial and non-financial features leads to a more accurate prediction of financial distress than exclusive examination of either financial or non-financial features. Our study bears implications for both investors and governmental regulators. Investors will be able to obtain a better understanding of the roles quantitative and qualitative features play in predicting corporate business crisis. Government regulators might be able to detect and prevent potential financial crises in early stage. Second, support vector machine, a relatively new learning method, is adopted to predict business crisis based on both financial and non-financial features. Our study integrates the non-financial features based on the concept of corporate governance to diagnose the financial health of a business. For enhancing the model's performance, feature selection is undertaken by employing stepwise regression to identify the critical features as the input variables.

We make several contributions to the literature. First, we document that effective corporate governance requires both internal and external measures, thereby enhancing the validity of the Cremers and Nair [11] findings. Second, we identify 42 corporate governance non-financial features that are related to firm value, an effort expected to significantly expand our knowledge of the internal governance factors linked to firm value beyond the sole (shareholder activism) variable suggested by Cremers and Nair [11].

The next section focuses on a theoretical overview of business crisis prediction. Section III introduces the proposed methods for business crisis prediction such as stepwise regression, genetic algorithm, multivariate statistical technology, and SVM. Section IV outlines the research experiment framework and design adopted by our study. The experiment results and discussion are presented in Section V. Finally, the conclusion is provided in Section VI.

II. THEORETICAL OVERVIEW

Business crisis prediction is not only an important but also a challenging problem stimulating numerous studies over the past decades. Early studies tend to treat financial ratios measuring profitability, liquidity and solvency as significant indicators for the detection of financial difficulties. However, reliance on these financial ratios can be problematic. The order of their importance, for example, remains unclear as different studies suggest different ratios as the major indicators of potential financial problems.

1. Financial Features and Financial Crises

The pioneering study of Beaver [2] introduces a univariate approach of discriminant analysis to predict financial distress. The method was later expanded into a multivariate framework by Altman [1]. Discriminant analysis had been the primary method of business failure prediction until 1980s during which the use of logistic regression method was emphasized. The standard discriminant analysis procedures assume that the variables used to characterize the members of the groups under investigation are in multivariate normal distribution. However, in real life, deviations from the normality assumptions are more likely to take place, and this violation may result in biased results. A non-linear logistic function is preferred over multivariate discriminant analysis (MDA), and there are researchers [1, 15, 16] claiming that even when all the assumptions of MDA hold, a logit model is virtually as efficient as a linear classifier. Considerable discrepancy is observed in the prediction accuracy reached by the three methods since using different methods leads to different prediction models that adopt different financial ratios.

Major financial features selected for financial distress prediction include financial leverage, long-term and short-term capital intensiveness, return on investment, EPS and debt coverage stability, etc. Selection of these features, however, is seldom based on a theory capable of explaining why and how certain financial factors are linked to corporate bankruptcy [15, 16]. We select variables using quantitative methods and carefully choose data sets from Taiwan's manufacturing industry. Despite the numerous definitions of business crises, the general meaning should include some narrower definitions like bankruptcy and shut-down and some broader definitions like failure, decline and distress. According to Beaver [2], a business crisis occurs when a firm announces its bankruptcy, bond default, over-drawn bank account or nonpayment of preferred stock dividends. As financial factors are mostly backward-looking, point-in-time measures, prediction models examining only financial features are inherently constrained. This paper accordingly would like to further explore the role of non-financial features in corporate business crisis prediction.

2. Non-Financial Features Related to Corporate Governance

According to the study by Günther and Grüning [15], 70 of the 145 surveyed German banks examine not only quantitative but also qualitative factors in credit risk assessment. Consideration of qualitative variables is found to help improve the percentage of companies correctly classified. While the eligibility of financial features as inputs for business crisis prediction is widely accepted, the role of non-financial features remains ambiguous. With financial scandals increasing in both frequency and size in these years, it becomes clear that the specific role of and interaction between different risk factors in financial scandals have to be analyzed in more details. These non-financial factors are usually selected based on experts' judgments and common business knowledge.

According to prior corporate governance literature [19, 20, 32], many listed companies in Taiwan still rely heavily on the support of their founding families to finance their operations, in marked contrast to companies in industrialized countries. In a sample of 141 companies listed on the Taiwan Stock Exchange (TSE), Claessens *et al*. [10] noted that 34% were family-controlled, where control was defined as having a 20% shareholding. If the criterion for control is reduced to a 10% shareholding, the percentage of family-controlled listed companies escalated to 47%. The percentage went on to hit 67.5% if the legal definition of insider shareholding is used. The extensive presence of family control in Taiwan's listed companies renders corporate governance a particularly crucial concern in financial distress prediction.

Existing studies on firms with a concentrated ownership structure, such as the one by Claessens *et al*. [9], primarily use the divergence between control and ownership as a measure of the agency conflict between majority and minority shareholders. However, the divergence measure can be difficult for investors to calculate accurately, especially when family-based controlling shareholders use pyramids and cross-holdings to leverage control or divert resources. A major conclusion of studies on companies with a concentrated ownership structure indicates that greater agency conflicts and weaker corporate governance are highly likely to exist when the majority of directors and all of the supervisors belong to a controlling family. Therefore, a firm's board structure can serve as an important indicator of whether the controlling family shareholder is committed to or entrenching corporate governance. On the other hand, concentrated ownership creates the conditions for a new agency problem because the interests of controlling and minority shareholders are not perfectly aligned, especially when there is a divergence between control and ownership. In such instances, corporate boards could play an important role in limiting the power of controlling shareholders to monitor important decisions [19].

Yeh and Woidtke [32] suggest that controlling shareholders entrench themselves by selecting both board members that are more likely to make decisions favoring their interests and those that are less likely to monitor when divergence goes up. Moreover, the resulting increase in board affiliation is associated with negative valuation in family-controlled firms. Recently corporate financial scandals in Taiwan betray a common feature consistent with the conclusion of related studies that larger agency conflicts and weaker corporate governance exist when the board is dominated by members closely affiliated with the controlling family.

In response to the extensive presence of concentrated ownership in Taiwan, we accordingly conduct regression model to select "Shareholding of Board Members-Current vs. Prior Year", "Ratio of Pledged Shares of Board Members", "Shareholding of Board Members", "Necessary Controlling Holding Shares", "Other Investment Assets" and "Board Member Bonus to Pretax Income" out of 42 original non-financial features as shown in Appendix A in our proposed financial distress prediction model.

III. USINESS CRISIS PREDICTION MODELS: THE BACKGROUND

Substantial literature can be found on business crisis prediction. We categorize the methods extensively used in prior research such as stepwise regression, genetic algorithm and multivariate statistical technology, etc. for corporate business crises prediction. Then, the SVM is briefly introduced.

1. Stepwise Regression Analysis

Model selection and parameter search play a crucial role in the performance of business crisis prediction. The stepwise selection identifies several variables as significant predictors. Prior researches indicate that the regression model has a better overall fit and a higher percentage of bankruptcy classification than the discriminant model [9, 10].

2. Genetic Algorithm

Genetic algorithms (GA) [22, 31] can be adopted to solve global optimization problems. The procedure starts with a set of randomly created or selected possible solutions, referred to as the population. Every individual in the population suggests a possible solution, referred to as a chromosome. Within every generation, a fitness function should be used to evaluate the quality of every chromosome to determine the probability of its surviving into the next generation; usually, the chromosomes with larger fitness have a higher survival probability. Thus, GA should select the chromosomes with larger fitness for reproduction by using operations like selection, crossover and mutation in order to form a new group of chromosomes which are more likely to reach the goal. This reproduction goes through one generation to another, until it converges on the individual generation with the most fitness for goal functions or the required number of generations is reached. The optimal solution is then determined [7].

Min, Lee, and Han [22] propose a genetic algorithm (GA) to search for the parameters of SVM for diagnosing business crisis; however, the model takes only finance features into consideration. Other features with substantially critical influence are not selected, and only the conventional binary GA is used [21]. Wu *et al*. [31] employ a real-valued genetic algorithm (GA) to optimize the parameters of SVM for predicting bankruptcy by using 19 financial variables. The real-valued genetic algorithm (RGA) uses a real value as a parameter of the chromosome in populations without performing the coding and encoding process before calculating the fitness values of individuals. Namely, RGA is more straightforward, faster and more efficient than other GA models such as binary genetic algorithm.

3. Multivariate Statistical Technology

Altman [1] introduces multivariate statistical technique known as discriminant analysis approach as an alternative to traditional ratio analysis for corporate bankruptcy prediction. He employs a sample of 66 corporations with 33 firms in each of the two groups with different asset sizes and reports Z-scores. He concludes that the model performs well with a 94% accuracy in predicting bankruptcy. He also claims that bankruptcy can be accurately predicted up to two years prior to actual failure with the accuracy diminishing rapidly after the second year [24]. Altman's Z-score model was brought to the attention of auditors via a 1974 article entitled "Evaluation of a Company as a Going-Concern." As a result, the updated model, or variations on it, has now been used by auditors and others to provide a bankruptcy risk signal for more than three decades. For example, the Altman model was adopted to examine prediction possibilities for the July 2002 WorldCom bankruptcy [8]. Grice and Ingram [33] reported that Altman's Z-score model declined when applied to various industries. In recently studies, several revised financial distress prediction models such as the revised Z-score models and the hybrid system [18, 29] have been demonstrated the results of highly adaptable in predicting bankruptcy.

4. SVM Model

As a relatively new algorithm in machine learning, support vector machine (SVM) was first developed by Boster, Guyon, and Vapnik [3] to provide better solutions to decision boundary than could be obtained using the traditional neural network. The machine learning techniques automatically extract knowledge from a data set and construct different model representations to explain the data set. The SVM approach has been put into several financial applications recently, mainly in the area of time series prediction and classification [26]. SVM belongs to the type of maximal margin classifier, in which the classification problem can be represented as an optimization process. Vapnik [30] showed how training a support vector machine for pattern recognition could lead to a quadratic optimization problem with bound constraints and one linear equality constraint. The basic procedure for applying SVM to a classification model can be summarized as follows [7]. First, the input vector is mapped into a feature space, which is possible with a higher dimension. The mapping is either linear or non-linear, depending on the kernel function selected. Then, within the feature space, the approach proceeds to seek an optimized division, i.e., to construct a hyper-plane that separates two or more classes. Using the structural risk minimization rule, the training of SVMs always seeks a globally optimized solution and avoids over-fitting. It has, therefore, the ability to deal with a large number of features. The decision function (or hyper-plane) determined by a SVM is composed of a set of support vectors selected from the training samples.

The SVM developed by Vapnik [30] implements the principle of Structural Risk Minimization by constructing an optimal separating hyper plane $w \cdot x + b = 0$. SVM uses a linear model to separate sample data through some nonlinear mapping from the input vectors into the high-dimensional feature space. Unlike most of the traditional neural network models which implement the Empirical Risk Minimization Principle, SVM seeks to minimize an upper bound of the generalization error rather than minimizing the training error. To make an efficient SVM model, two extra parameters: C and σ^2 (sigma squared) have to be carefully predetermined. The first parameter, C, determines the trade-offs between the minimization of the fitting error and the minimization of the model complexity. The second parameter is the bandwidth of the radial basic function (RBF) kernel. To find the optimal hyper plane $\{x \in S\ (w, x) + b = 0\}$, the norm of the vector w needs to be minimized while the margin between the two classes $1/||w||$ should be maximized.

$$\min_{i=1,\ldots n} \left|\left(w, x\right) + b\right| = 1 \tag{1}$$

Popular kernel functions in machine learning theories are summarized as follows. According to Lagrange multiplier, decision function is built as follows:

$$Q(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{ij=1}^{l} \alpha_i \alpha_j y_i y_j K\left(x_i, x_j\right) \tag{2}$$

subject to $0 \le \alpha_i \le C$,

$$\sum_{i=1}^{l} \alpha_i y_i = 0$$

with the decision function $f(x) = \text{sign}\left(\sum_{i=1}^{l} y_i \alpha_i k\left(x, x_i\right) + b\right)$

SVM works as a maximal margin classifier in which the classification problem can be represented as an optimization process. Support vectors are a subset of training data used to define the boundary between two classes. As suggested by Vapnik [30], SVM can be generalized well even in high-dimensional spaces under small training sample conditions, indicating a learning ability independent of the feature space dimensionality.

The training of SVMs is equivalent to solving a linearly constrained quadratic programming, helping reach a solution that is unique, optimal and absent from local minima. It is robust to outliers. It reduces the effect of outliers by using the margin parameter C to control the misclassification error. Moreover, with Vapnik's e-insensitive loss function, SVM can model nonlinear functional relationships difficult to be modeled by other techniques [30].

These characteristics make SVM a strong candidate in predicting financial distress. Therefore, our proposed model defines the bankruptcy problem as a nonlinear problem and uses the RBF kernel below to optimize the hyper plan.

$$\text{(RBF): } K\ (x, y) = e^{-\|x-y\|2/2\sigma^2} \tag{3}$$

In (3), σ^2 denotes the variance of the Gaussian kernel.

The major difference between traditional statistical methods and machine learning methods is that statistical methods usually require the researchers to impose structures onto different models, such as the linearity in the multiple regression analysis, and to construct the model by estimating parameters to fit the data or observation, while machine learning techniques allow learning the particular structure of the model from the data [16].

Prior researches on bankruptcy prediction have pinpointed a considerable number of significant predictors of business failure [2, 14]. In previous studies, a comprehensive list of financial ratios has been developed and grouped into the eight categories of profitability, liquidity, solvency, degree of economic distress, leverage, efficiency, variability, and time.

Studies on corporate boards of directors are generally restricted to large firms in US where investor protection is strong and ownership is disperse and tend to treat board composition as being exogenous [32]. Corporate governance is therefore seldom taken into consideration as a contributing factor in corporate financial distress. However, studies focusing on emerging markets indicate that corporate governance can be a significant issue as ownership structures tend to be concentrated in most countries outside the US. Therefore, the non-financial features we select most evolve the issue of corporate governance.

IV. EXPERIMENT FRAMEWORK AND DESIGN

In this section, we present the experiment framework and design of our proposed model. A publicly listed firm is regarded to encounter business crisis and turns into a distressed company when declared for full-value delivery, stock transaction suspension, re-construction, bankruptcy or withdrawal from the stock market. Based on the above criteria, 54 distressed and 54 non-distressed (as matched sample) companies are identified in Taiwan during the period from 2001 to 2005 according to Taiwan Economic Journal (TEJ) databank that incorporates two extra criteria: 1. The sampled firms should have at least four quarters of complete public information before the business crisis happens; 2. There should be sufficient comparable companies with similar size and in the same industry to serve as contrary samples. In general, business crises could be classified into two types. The first type refers to the scenario in which a given business entity after several professionals' independent evaluations is consistently recognized as lacking the capital for business management; major predictors of this type of business crisis are mainly financial in nature and include current ratio, quick ratio, debt ratio, receivables turnover ratio, and fixed asset turnover ratio [14, 25]. The second type refers to the situation when a firm with stock released on the public market is declared for full-value delivery or legally put in transaction suspension, re-construction, bankruptcy or withdrawal from the stock market. Indicators of this type of business crisis usually move beyond conventional financial information to touch upon non-financial features such as the factors of family holding shares, necessary controlling shares, frequently in board of director and manager changes, and stockholder's behaviors.

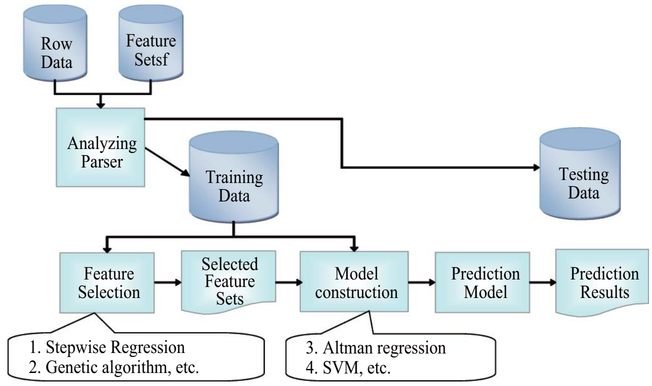

Fig 1. Overall procedure of modeling.

Feature selection can adopt stepwise regression, genetic algorithm, etc, while model construction can utilize the methods such as multivariate statistical technique, SVM and so on. Figure 1 illustrates the overall procedure of modeling the business crisis prediction as we have described in Section III.

1. The Experiment Design and Tools

In our proposed regression-SVM model, the SVM parameters are dynamically optimized by implementing the stepwise regression process. After a survey on the features recommended by scholars and their availability, stepwise regression using SPSS 13.0 [27] was performed to select features for the proposed prediction model, and a level lower than 5% is considered statistically significant. An Analyzing Parser is developed to process the financial statements retrieved from TEJ (Taiwan Economic Journal) databank. We use the Analyzing Parser to create both financial and non-financial features. These data are used either as training data to construct the prediction model or as the testing data to validate the proposed model through SVM by using these optimal values. In general, the radial basis function (RBF) is suggested for SVM. The RBF kernel nonlinearly maps the samples into the high-dimensional space, so it can handle nonlinear problems. We use LIBSVM software [6] to construct the classification model and choose RBF as the kernel function. Since the performance is generally evaluated by cost, e.g. classification accuracy or mean square error (MSE), we also try to change the values of "gamma" and "cost" in order to enhance prediction results. Namely, the stepwise regression tries to search the optimal values to enable SVM to fit various datasets.

The holdout method, sometimes called test sample estimation, partitions the data into two mutually exclusive subsets called a training set and a testing set, or a holdout set. Generally, about two thirds of the data are used as the training set and the remaining one third as the testing set. The training set is given to the inducer, and the induced classifier is tested on the test set. The comparison is based on a training set with equal proportion of distressed or non-distressed firms. The testing data consists of both distressed and non-distressed companies. It is important to note that the training and testing sets are mutually exclusive.

The objective of this research is to investigate if the incorporation of non-financial features [10, 19], such as pledged shares of board members, change in stock ownership of board members, and frequent CPA change, help increase financial distress prediction quality in addition to the traditional focus on financial information. Each of the steps is summarized as follows:

1. Stepwise regression is applied and SPSS 13.0 used to select the features for our new model;
2. Initial population is randomized.
3. An Analyzer Parser, is developed to code the features, such as the common ratios, and to create training data based on the features determined in Step 1 and 2;
4. The training data are fed into the SVM tool to create the prediction models for our experiment.
5. Finally, the testing data are prepared using the Analyzing Parser in a manner similar to the one for training data in Step 3, and the prediction results are obtained by applying the prediction models from Step 3.

2. Feature Selection

To launch experiments with our new model, we first survey literature related to corporate governance [19, 20] and analyze the distressed firms in Taiwan to select the variables which indicate significant differences between the distressed group and non-distressed group. Then, the final input features were selected through stepwise logistic regression analysis and correlation analysis.

The SVM rests on the data generated from the year-end financial statements of the firms and is carried out to identify the most important predictors in bankruptcy classification. Based on the outcome of the stepwise selection, ten variables are identified as significant predictors as shown in Table 1, which includes 4 financial features (out of 23) and 6 non-financial features (out of 42) related to corporate governance. All the financial and non-financial features considered in this study are listed in Appendix A. As mentioned before, every feature should include at least 4 quarters of data before the business crisis. The input variables of all the financial features in all models are the same. The bootstrap technique has been widely used in financial research to evaluate the external validity of model in prediction.

In this study, the sample covers 54 publicly traded firms encountering financial crises during the period from 2001 to 2005 in Taiwan while their non-distressed counterparts (54 firms in total) with a similar size and in the same industry are also surveyed. The distressed firms are selected based on the quarterly financial reports of listed companies in Taiwan collected in the TEJ databank. We gather 312 (54 * 2 * 3 = 312) observations from the 3-year annual reports of the 108 sampled firms.

Table 1. The features for business crises prediction.

Features	Definition
Financial features	
F1	Debt Ratio
F2	Accounts Receivable Turnover Ratio
F3	Assets Turnover
F4	Operating Income to Capital
Non-Financial features	
N1	Shareholding of Board Members - Current vs. Prior Year
N2	Ratio of Pledged Shares of Board Members
N3	Shareholding of Board Members
N4	Necessary Controlling Shares
N5	Other Investment Assets
N6	Board Member Bonus to Pretax Income

Table 2. Profile analysis – means and standard deviations by feature.

Firm types	Distressed firms		Non-distressed firms	
Features	Mean	Std. Dev.	Mean	Std. Dev.
F1	0.684	0.160	0.405	0.130
F2	6.336	6.974	15.45	34.81
F3	0.547	0.441	0.812	0.537
F4	-10.98	17.12	8.837	9.252
N1	-2.873	5.426	-0.791	3.798
N2	44.98	36.27	17.29	25.96
N3	13.84	10.62	22.82	13.14
N4	10.72	6.433	12.48	6.093
N5	14.87	12.43	23.95	16.14
N6	0.0	0.0	1.034	1.150

Besides, Type I and Type II errors are analyzed in these experiments. Type I error occurs when a firm is predicted to be healthy but is in fact distressed; Type II error, on the other hand, takes place when a firm is predicted to be distressed but is in fact healthy.

A Summary of profile analysis by features is shown in Table 2. We have utilized "exhausted search" method to process all the experiments. For each experiment, SVM is used to predict business crisis for the sampled companies, and the prediction ability of the proposed model is evaluated, which has shown good performance in model selection. When performing the cross-validation procedure for SVM, we choose the leave-one-out sampling approach due to the size of our sample data.

V. EXPERIMENT RESULTS AND DISCUSSION

1. Performance Comparison

For performance comparison, we create three different prediction models: Model 1 is based exclusively on our selected financial features; Model 2 is based solely on non-financial features related to corporate governance; and finally, Model 3, the proposed hybrid model, combines both financial and non-financial features. Different types of errors result in different penalty costs. As presented earlier, 54 distressed firms in the years of 2001-2005 are analyzed against 54 non-distressed counterparts. We first compare the prediction accuracy of the three models using the data, both financial and non-financial, one year prior to the bankruptcy of each distressed firm. This prediction is also known as the 1-year-ahead forecast [12]. For benchmark purpose, we also apply the Z-score model with the same features used in the three models. In addition to the 1-year-ahead forecast, we extend our analysis to cover three consecutive years of financial statements for each of the studied 108 firms in order to examine the longer term prediction power of each of the three models. In other words, our studies perform three forecasts: 1-year-ahead, 2-year-ahead, and 3-year-ahead.

In Model 1, we endeavor to examine the financial model known for its capability to solve classification problems in financial prediction so as to launch a comparison with our new model. Based on the best experiment on Model 1, F1 (Debt Ratio), F2 (Accounts Receivable Turnover Ratio), F3 (Assets Turnover) and F4 (Operating Income to Capital) emerge to be the more accurate of all the 23 financial predictors listed in Appendix A. The average accuracy of the 1-year-ahead forecast is 88.89% with Type I and Type II error rates being 12.96% and 9.26%, respectively. Type I error (misclassifying a distressed firm as a healthy one) appears more frequently than Type II error (misclassifying a healthy firm as a distressed one). These results are summarized in Table 3.

Model 2 examines non-financial features to predict distressed firms with SVM. According to the results of the experiment on Model 2, N1 (Shareholding of Board Members – Current vs. Prior year), N2 (Ratio of Pledged Shares of Board Members), N4 (Necessary Controlling Holding Shares), N5 (Other Investment Assets) and N6 (Board Member Bonus to Pretax Income), appear to be the more accurate of all the non-financial predictors covered in Model 2. As summarized in Table 3, the average accuracy of the 1-year-ahead forecast in Model 2 is 87.96% with a Type I error rate of 5.56% and a Type II error rate of 18.52%. Compared to Model 1, Model 2 sustains an improved prediction performance thanks to its lower rate of Type I error. The prediction capability of various models for longer terms is discussed later.

For Model 3, the proposed hybrid Model, F1 (Debt Ratio), F2 (Accounts Receivable Turnover Ratio), F4 (Operating Income to Capital), N1 (Shareholding of Board Members-Current vs. Prior Year), N2 (Ratio of Pledged Shares of Board Members), N5 (Other Investment Assets) and N6 (Board Member Bonus to Pretax Income) are identified as the more accurate of all the adopted financial and non-financial features. The average accuracy for the 1-year-ahead forecast reads 94.44%, significantly superior to those of Model 1

Table 3. Financial and non-financial model comparison with SVM.

Evaluation criterion	Financial (Model 1)	Non-financial (Model 2)	Hybrid (Model 3)
Type I error	0.1296	0.0556	0.0556
Type II error	0.0926	0.1852	0.0556
Brier Score (BS)	0.1111	0.1204	0.0556
Average accuracy	0.8889	0.8796	0.9444
Feature selected	[F1], [F2], [F3], [F4]	[N1], [N2], [N4], [N5], [N6]	[F1], [F2], [F4], [N1], [N2], [N5], [N6]

*the experiment using cross-validation

(88.89%) and Model 2 (87.96%). Model 3 also performs better than the other two models in terms of Type I and Type II errors as both reports a rate of 5.56%. Compared with the other two models, Type I error and Type II error occurs with a less frequency in Model 3. In actual practice, the cost of misclassifying a failed firm into a healthy one (Type I error) is likely to be much greater than that of misclassifying a healthy firm into a failed one (Type II error). As indicated above, the Type I and Type II errors of Model 3 were much lower than those of Model 1 and Model 2. Empirical results indicate that Model 3 examining both financial and non-financial features can serve as a promising alternative for existing financial distress prediction models.

We further adopted Brier Score (BS) [4] for comparison of prediction accuracy. The Brier Score (BS) is a measure of prediction accuracy well-known in meteorology and medical science. It is formulated as $[\mathrm{BS} = \frac{1}{\mathrm{n}}\sum_{i}^{n}(\theta_i - p_i)^2 1)]$ where θ_i is a binary indicator for the actual realization of the default variable (1 if default, 0 if no default) and p_i, is the estimated probability of default. The difference between the Brier Score and the percentage of correctly classified observations is that the former is more sensitive to the level of the estimated probabilities. The Brier Score takes the estimated probabilities directly into account. According to the results presented in Table 3, the combination of financial and non-financial features achieves a lower average Brier Score (BS) of 5.56% after taking into consideration of all experiment results.

As Table 3 shows, the average accuracy for 1-year-ahead forecast of all three models falls in the range between 87.96% and 94.44%. The proposed hybrid model is able to predict bankruptcy one year ahead with an impressive accuracy of 94.44%. Compared with Model 1, Model 3 takes non-financial features into account and leads to an increase in average accuracy from 88.89% to 94.44%. This implies that non-financial features, especially those vulnerable to the manipulation of a firm's board members, deserve equal scrutiny in predicting financial distress. Therefore, combined consideration of both financial and non-financial features can be expected to greatly

Table 4. Prediction accuracies (Z: Z-score models, M: support vector machines.).

Financial Models		Non-financial Models		Hybrid Models	
Z_1 (%)	M_1 (%)	Z_2 (%)	M_2 (%)	Z_3 (%)	M_3 (%)
85.18	88.89	81.48	87.96	90.74	94.44

enhance the accuracy of a financial distress prediction model.

In summary, a hybrid model encompassing both financial and non-financial features can be expected to achieve a more accurate prediction of corporate financial distress than a model based exclusively on either non-financial or financial feathers.

2. Comparison of SVM with Z-score Model

Or benchmark purpose, we apply Z-score model to construct three models Z_1, Z_2, and Z_3 as their SVM counterparts M_1 (Model 1), M_2 (Model 2) and M_3 (Model 3). The prediction accuracies of the 1-year-ahead forecast are shown in Table 4, where the Z-score models consistently fall short of their SVM counterpart models. For example, Z_3 yields a 90.74% accuracy that is lower than the one achieved by M_3. Furthermore, in terms of prediction accuracy, the Z-score models report a similar trend as the SVM models as shown in Fig. 2, namely, Z_3 outperforms both Z_1 and Z_2 as M_3 outperforms both M_1 and M_2. Therefore, we conclude that the hybrid model, either Z_3 or M_3, appears to be the best model in prediction accuracy among the three models, whereas, the non-financial model, either Z_2 or M_2, seems to be the least desirable model.

3. The Analysis of Predictive Accuracy for Longer-Term Forecast

We further conduct additional experiment to observe the effect of the prediction capability of these models for longer terms. Table 5 shows the results of applying the three models for 2-year-ahead forecast and 3-year-ahead prediction. Model 1 sustains an accuracy of 78.7% for 2-year-ahead forecast and 75.92% for 3-year-ahead forecast. The accuracies for 2-year-ahead and 3-year-ahead forecasts read respectively 70.37% and 71.29% for Model 2 and 75.93%, and 74.07% for Model 3. As the results indicate, for longer-term forecast, Model 1 takes the lead in terms of predictive accuracy, followed respectively by Model 3 and Model 2. In general, the financial condition of TSE listed companies can be better predicted using the SVM model for long-term forecasts since the prediction accuracy of SVM Model 1 is slightly higher than Model 3 and Model 1. However, Model 1 focuses only on financial ratios related to a firm's business performance whereas the proposed model adds on the top of financial ratios several non-financial features concerning the behaviors of a company's board members. Detailed analysis of cases indicates that a distressed firm is more likely to engage in manipulations related to the selected non-financial features as the firm approaches nearer to the verge of bankruptcy [31, 32]. In other words, manipulation by a firm's board members tend to occur when the firm's financial distress proves to be imminent or unavoidable; one year before of a firm's bankruptcy is therefore a better or more urgent timing than two or three years ahead. This may explain the greater predictive accuracy of Model 3, our proposed hybrid model, in 1-year-ahead forecast and its relatively lower accuracy in 2- and 3-year-ahead forecasts when compared to the financial-feature-only Model 1. However, when it comes to the average of the three (1~3-year-head) forecasts, Model 3 remains in the first place with an average predictive accuracy of 81.4%, followed respectively by Model 1 (81.1%) and Model 2 (76.5%) and suggesting that, in general, the financial status of listed companies in Taiwan can be better predicted using our proposed SVM-based hybrid model.

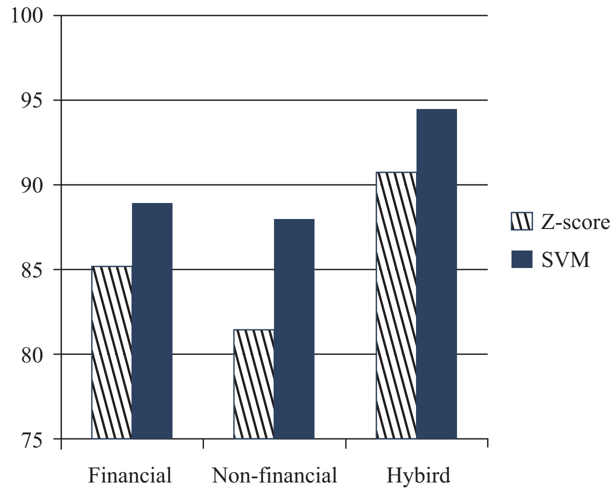

Fig 2. Comparison of SVM models with Z-score models.

Table 5. The 1-year ahead to 3-year ahead forecasts of Model 1, 2, and 3.

	Model 1	Model 2	Model 3
1-year-ahead forecast	88.89%	87.96%	94.44%
2-year-ahead forecast	78.70%	70.37%	75.93%
3-year-ahead forecast	75.92%	71.29%	74.07%

VI. CONCLUSION

This paper proposes a model based on support vector machine and taking into consideration of both financial and non-financial features for business crisis prediction. As the extensive presence of concentrated ownership in the public listed companies in Taiwan. We analyze via the SVM method several non-financial features related to corporate governance, notably "Shareholding of Board Members-Current vs. Prior Year", "Ratio of Pledged Shares of Board Members", "Shareholding of Board Members", "Necessary Controlling Holding Shares", "Other Investment Assets", and "Board Member

Bonus to Pretax Income". The empirical results indicate that examining the selected non-financial features in addition to traditional financial indictors provides a promising solution for assessing the risk of corporate bankruptcy. The hybrid model is capable of achieving an improved predictive accuracy, especially for one-year-ahead forecast. The overall predictive accuracy rate achieved by our proposed model for business crisis prediction reads 94.4%, superior to those of the one based exclusively on financial ratios and the one considering only non-financial feature. Inclusion of these non-financial features appears to enhance the performance of business crisis prediction. These non-financial features related to corporate governance may merit consideration in future researches, especially those focusing on emerging markets populated with firms characterized by concentrated ownership.

Moreover, for benchmark purpose, we also compare the SVM models with Z-score model and detect in the Z-score models in a similar trend which the proposed hybrid model outperforms both the financial-only and nonfinancial-only models in terms of predictive accuracy. In addition to the same pattern, the SVM models outperform the Z-score models. Therefore, we conclude that the hybrid model of SVM appears to be the best model in predictive accuracy among the three models, whereas, the non-financial model seems to be least desirable model.

There are, on the other hand, limitations in this article that call for further researches. Our models are inevitably affected by several factors. First of all, the predictive accuracy might be further improved in the future by considering to pair sampled companies by industry or to extend the survey period. It should further be noted that in reaction against the recent outbreak of corporate financial scandals in Taiwan and overseas, we have paid special attention to the roles of ownership structure and corporate governance in business crisis prediction. Selection of non-financial features is therefore based on attributes related to corporate governance. This exclusive focus on corporate governance-related factors has prevented us from considering in our present study other potentially influential non-financial features, such as market share, management style, and industry prospect. Further researches may be conducted to explore such potential non-financial indicators.

APPENDIX A

Table A. A list of financial features.

No.	Financial Features
1.	Debt Ratio
2.	Long-term Capital to Fixed Assets Ratio
3.	Long-term Capital to Fixed Assets and Long-term Equity Ratio
4.	Current Liability to Total Assets
5.	Current Ratio
6.	Quick Ratio
7.	Time Interest Earned
8.	Working Capital to Total Assets
9.	Accounts Receivable Turnover Ratio
10.	Average Number of Days Accounts Receivable
11.	Inventory Turnover
12.	Days Sales in Inventory
13.	Average Number of Days Accounts Payable Outstanding
14.	Fixed Assets Turnover
15.	Assets Turnover
16.	Return on Assets
17.	Operating Income to Capital
18.	Earnings before Income Tax to Capital
19.	Income to Capital
20.	Earnings Per Share
21.	Sales Per Employee
22.	Operating Income Per Employee
23.	Long-term Investment to Assets

Table B. A list of non-financial features.

No.	Non-Financial Features
1.	Director & Supervisor Holding Shares
2.	Director & Supervisor Holding Shares-Current-Prior year
3.	Pledged Share Ratio of Director & Supervisor
4.	Manager Holding Shares Ratio
5.	Director & Supervisor Holding Shares
6.	Director Holding Shares
7.	Supervisor Holding Shares
8.	Director & Supervisor Pledged Shares
9.	Director Pledged Shares
10.	Supervisor Pledged Shares
11.	Main Shareholders' Holding Shares
12.	Family Individual Holding Shares
13.	Family Unlisted Company Holding Shares
14.	Family Funding Holding Shares
15.	Family Listed Company Holding Shares
16.	Manager Holding Shares
17.	Outside Individual Holding Shares
18.	Outside Unlisted Company Holding Shares
19.	Outside Funding Holding Shares
20.	Outside Listed Company Holding Shares
21.	Controlling Holding Shares
22.	Direct Holding Shares
23.	Earnings Appropriation
24.	Necessary Controlling Holding Shares
25.	Excess Holding Shares
26.	Other Investment to Equity
27.	Other Investment to Assets
28.	Director Bonus
29	Employee Bonus to PreTax Income
30.	Employee's Cash Bouns
31.	Employee's Stock Bonus
32.	Sales-Related Party
33.	Purchases-Related Party
34.	Finance-Related Party
35.	Disposal Gain (Loss) – Related Party
36.	Frequent in Board of Director Change
37.	Frequent in General Manager Change
38.	Frequent in CFO Change
39.	Director & Supervisor Bouns to Pertax income
40.	Average Bonus per Director & Supervisor
41.	Disposal Gain (Loss) – Related Party
42.	Frequent in Board of Director Change

REFERENCES

1. Altman, E. I., "Financial ratio, discriminant analysis, and the prediction of corporate bankruptcy," *Journal of Finance*, Vol. 23, pp. 589-609, September (1968).
2. Beaver, W., "Financial ratios as predictors of failure. Empirical research in accounting: selected studies," *Journal of Accounting Research*, Vol. 4, pp. 71-111 (1966).
3. Boster, B., Guyon, I., and Vapnik, V. N., "A training algorithm for optimal margin classifiers," *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, pp. 144-152 (1992).
4. Brier, G. W., "Verification of forecasts expressed in terms of probability," *Monthly Weather Review*, Vol. 78, pp. 1-3 (1950).
5. Burbidge, N. M., *Organ Allocation and Patient Responsibility: Reexamining the Concept of Responsibility in Light of the Thought of Emmanuel Levinas*, Ph. D. Dissertation, University of Ottawa, Canada, 337 pages; AAT NQ66129 (2001).
6. Chang, C. C. and Lin, C. J., "LIBSVM: A library for support vector machines," Software available at http://www.csie.ntu.edu.tw/cjlin/libsvm. (2001).
7. Chen, L. and Hsiao, H., "Feature selection to diagnose a business crisis by using a real GA-based support vector machine: An empirical study," *Expert Systems with Applications*, Vol. 30 (2007).
8. Chuvakhin, N. and Gertmenian, L. W., "Predicting bankruptcy in the WorldCom age," *Journal of Contemporary Business Practice 6* (1). Available at: http://gbr.pepperdine.edu/031/bankuptcy.html (2003).
9. Claessens, S., Djankov, S., Fan, J. P. H. and Lang, L. H. P., ''Disentangling the incentive and entrenchment effects of large shareholdings,'' *The Journal of Finance*, Vol. 57, No. 6, p. 2741-2771 (2002).
10. Claessens, S., Djankov, S., and Xu, L. C., ''Corporate performance in the East Asian financial crisis," *The World Bank Research Observer*, Vol. 15, No. 1, pp. 23-46 (2000).
11. Cremers, K. J. M. and Nair, V. B., "Governance mechanisms and equity prices," *Journal of Finance*, Vol. 60, pp. 2859-2894 (2005).
12. Ding, Y., Song, X., and Zen, Y., "Forecasting financial condition of Chinese listed companies based on support vector machine," *Expert Systems with Applications*, Vol. 34, pp. 3081-3089 (2008).
13. Eichengreen, B.,''Is greater private sector burden sharing impossible?,'' *Finance & Development*, Washington: Sep. Vol. 36, No. 3, pp. 16-19, (1999).
14. Gilson, S. C., ''Management turnover and financial distress,'' *Journal of Financial Economics*, Vol. 25, No. 2, pp. 241-262 (1989).
15. Günther, T. and Grüning, M., "Einsatz von Insolvenzprognoseverfahren bei der Kreditwürdigkeitsprüfung im Firmenkundenbereich," *Die Betriebswirtschaft*, Vol. 60, pp. 39-59 (2000).
16. Huang, Z., Chen, H., Hsu, C. J., Chen, W. H., and Wu, S., "Credit rating analysis with support vector machines and neural networks: a market comparative study," *Decision Support Systems*, Vol. 37, pp. 543-558 (2004)
17. http://www.kpmg.com/xbrl/XBRL_Home.asp
18. Lee, K. C., Han, I., and Kwon, Y., "Hybrid neural network models for bankruptcy predictions," *Decision Support Systems*, Vol. 18, No. 1, pp. 63-72 (1996).
19. Lee, T. S., Woidtke, T., and Yeh, Y. H., "Corporate governance and financial distress: Evidence from Taiwan," *Corporate Governance*, Vol. 12, pp. 378-388 (2004).
20. Lien, Y., Piesse, J., Strange, R., and Filatotchev, I., "The role of corporate governance in FDI decisions: Evidence from Taiwan," *International Business Review*, Vol. 14, pp. 739-763 (2005).
21. Mao, H., Carson, J. M., Ostaszewski, K. M., and Luo, S., ''Pricing life insurance: Combining economic, financial, and actuarial approaches," *Journal of Insurance Issues*, Vol. 27, No. 2, pp. 134-159 (2004).
22. Min, S. H., Lee, J., and Han, I., "Hybrid genetic algorithms and support vector machines for bankruptcy prediction," *Expert Systems with Applications*, Vol. 31, pp. 652-660 (2006).
23. Ohlson, J. A., "Financial ratios and the probabilistic prediction of bankruptcy," *Journal of Accounting Research*, Vol. 18, pp. 109-31, Spring (1980).
24. Ozkan-Gunay, E. N. and Ozkan, M., "Prediction of bank failures in emerging financial markets: An ANN approach," *The Journal of Risk Finance*, Vol. 8, No. 5, pp. 465-480 (2007).
25. Sharmaand, S. and Mahajan, V., "Early warning indicators of business failure," *Journal of Marketing*, fall, pp. 80-89 (1980).
26. Shin, K. S., Lee, T. S., and Kim, H. J., "An application of support vector machines in bankruptcy prediction model," *Expert Systems with Application*, Vol. 28, pp. 127-135 (2005).
27. SPSS 13.0, url://www.spss.com
28. Taffler, R. and Tisshaw, H., "Going, going, gone-four factors which predict," *Accountancy*, Vol. 88, pp. 50-54, March (1977).
29. Tam, K. Y. and Kiang, M. Y., "Managerial applications of neural networks: The case of bank failure predictions," *Management Science*, Vol. 38, No. 7, 926-947 (1992).
30. Vapnik, V., *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, NY (1995).
31. Wu, C. H., Tzeng, G. H., Goo, Y. J., and Fang, W. C., "A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy," *Expert Systems with Applications*, Vol. 32, pp. 397-408 (2007).
32. Yeh, Y. H. and Woidtke, T., "Commitment or entrenchment? Controlling shareholders and board composition," *Journal of Banking & Finance*, Vol. 29, pp. 1857-1885 (2005).