The Role of Non-financial Features in Business Crisis Prediction

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

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

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 for predicting bankruptcy. Financial data have been widely used by researchers to predict business crisis, but few studies exploit the use of non-financial indicators in corporate governance to construct business crisis prediction model. This article introduces a prediction model based on a relatively new machine learning technique, support vector machines (SVM) for the field of business crisis prediction. 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.

INTRODUCTION

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 much broader scope. Eichengreen (1999) [11] identifies the policies of the new international financial architecture as crisis prevention, crisis prediction and crisis management. Financial indicators have been extensively consulted to predict financial crises by former researchers. The most common business crisis prediction methodologies are financial ratio and peer group analysis, comprehensive risk assessment systems, and statistical and econometric models [22].

Yeh and Woidtke (2005) [29] 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 frequent changes in Certified Public Accountant (CPA) by distressed companies prior to bankruptcy. We have therefore included the non-financial features such as corporate governance and CPA change factor 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. The former include linear regression, linear multivariate discriminant analysis (MDA), logit analysis and multidimensional scaling while the latter consist 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 (1992) [3] has been successfully applied in many areas, including financial time series forecasting, credit scoring, and drug design [5]. 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 search for predictors to help users identify underlying characteristics of distressed firms.

The aim of this paper is twofold. First, this paper not only explores the role of financial feature but also the role of 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 separate examination of either financial or non-financial features. Our study bears implications for both investors and governmental regulators. Investors will

Author for Correspondence: Fengyi Lin. E-mail: fengyi@ntut.edu.tw
**Associate Professor, Department of Business Management, National Taipei University of Technology*
***Professor, Department of Computer Science and Engineering, National Taiwan Ocean University*
****Assistant Professor, Department and Graduate Institute of Finance, National Yulin University*
*****Master Student, Department of Computer Science and Engineering, National Taiwan Ocean University*

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, were adopted to predict business crisis based on both financial and non-financial feature. 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.

The next section focuses on a theoretical overview of business crisis prediction. Section 3 introduces the proposed methods for business crisis prediction such as stepwise regression, genetic algorithm, multivariate statistical technology, SVM, etc. Section 4 outlines the research experiment framework and design adopted by our study. The experiment results are presented and discussed in Section 5. Finally, the conclusion is provided in Section 6.

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 like 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.

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 (1968) [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 prior researchers [1,14,15] claim 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. Despite the numerous definitions of business crises, the general meaning should include some narrow 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 would like to further explore the role of non-financial features in corporate business crisis prediction.

NON-FINANCIAL FEATURES RELATED TO CORPORATE GOVERNANCE

According to the study by Günther and Grüning (2000) [13], 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. Whereas 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 [17,18,29], 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 makes corporate governance a particularly crucial concern in financial distress prediction.

Existing studies on firms with a concentrated ownership structure, such as 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 [17].

Yeh and Woidtke (2005) [29] suggest that controlling shareholders entrench themselves further 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 corporate Taiwan, we accordingly select pledged shares of board members, real stock ownership of board members, and change in stock ownership of board members as the non-financial features in our proposed financial distress prediction model. Moreover, as several distressed firms, notably Rebar and Procomp Informatics, were found to change their auditors frequently before they went into bankruptcy. Thus, the frequency of changing CPA has also been included as one of our non-financial features.

BUSINESS CRISIS PREDICTION MODELS: THE BACKGROUND

Substantial literature can be found in business crisis prediction history. We categorized 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.

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].

GENETIC ALGORITHM

Genetic algorithms (GA) [20, 28] 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 (2006) [20] 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 [19]. Wu *et al.* (2007) [28] 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 coding and encoding process before calculates the fitness values of individuals. Namely, RGA is more straightforward, faster and more efficient than other GA models such as binary genetic algorithm.

MULTIVARIATE STATISTICAL TECHNOLOGY

Altman (1968) [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 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 [22]. Altman's Z-score model was brought to the attention of auditors via a 1974 article titled "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]. In recently studies, several revised financial distressed models such as the revised the Z score and ZETA models and the hybrid system [16, 27] have been demonstrated the results of highly adaptable and outperformed in predicting bankruptcy.

SVM MODEL

New algorithms in machine learning, Support vector machine (SVM), was developed by Boster, Guyon, and Vapnik (1992) [3] to provide better solutions to decision boundary than could be obtained using the traditional neural network. It is based on the Structural Risk Minimization principle from computational learning theory. 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 [24]. SVM belongs to the type of maximal margin classifier, in which the classification problem can be represented as an optimization process. Vapnik 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, map the input vector 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, seek an optimized division, i.e., 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 major difference between traditional statistical methods and machine learning methods is that statistical methods usually need the researchers to impose structures to 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 also allow learning the particular structure of the model from the data [14].

Prior researches on bankruptcy prediction have pinpointed a considerable number of significant predictors of business failure [2,12]. In previous studies, a comprehensive list of financial ratios has been developed and grouped into the following eight categories of profitability, liquidity, solvency, degree of economic distress, leverage, efficiency, variability, and time.

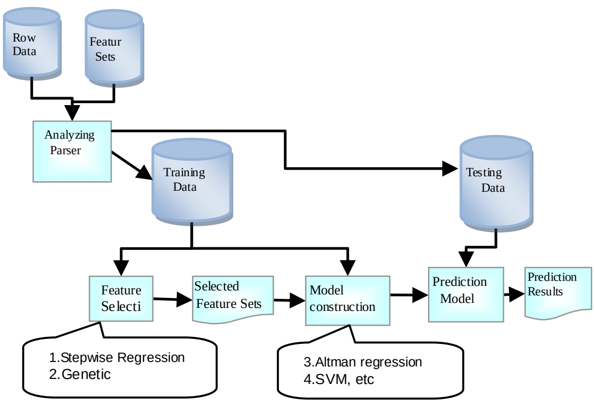
Studies on corporate boards of directors are generally restricted to large firms in US where investor protection is strong and ownership is disperse and tend to treat board composition as being exogenous [29]. Corporate governance is therefore seldom taken into consideration as a contributing factor 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 all evolve around the issue of corporate governance. For example, pledged shares of board members, real stock ownership of board members, and change in stock ownership of board members may exert direct and drastic impacts on business crisis prediction.

EXPERIMENT FRAMEWORK AND DESIGN

In this section, we present the environment and the tools, the experiment design, and the experiment results 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, and bankruptcy or goes out-of-market. Based on the above criteria, 26 distressed and 26 non-distressed (as matched sample) companies are identified in Taiwan during the period from 2003 to 2007 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, liability ratio, receivables turnover, cash flow, and total asset turnover [12,23] 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 corporate governance, frequently CPA changes, and stockholder's behaviors.

Feature selection can adopt stepwise regression, genetic algorithm, etc, while model construction can utilize the methods such as multivariate statistical technology, SVM and so on. Figure 1 illustrates the overall procedure of modeling the business crisis prediction as we have described in Section 3.

Figure 1. Overall procedure of modeling



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 [25] was then performed to select features for the proposed prediction model and a level lower than 5% is considered statistically significant. The Analyzing Parser is developed to process the financial statements retrieved from TEJ (Taiwan Economic Journal) databank. We use 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 test set, or a holdout set. About two thirds of the data are commonly used as the training set and the remaining one third are then used as the test 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, 17], 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. A *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.

FEATURE SELECTION

To launch experiments with our new model, we first survey literature related to corporate governance

[17, 18] and analyze the distressed firms in Taiwan to select the variables which indicate significant differences between the distressed group and non-distressed group. Then we select the final input features 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, eight variables are identified as significant predictors, including 3 financial features and 5 non-financial attributes related to corporate governance. 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.

Table 1: The features of business crises

Features	Meaning
Financial feature	
F1	Long-term investment/total assets
F2	Total liabilities/total assets
F3	Positive income statement with negative operating cash flow
Non-Financial feature	
N1	Pledged shares of board members
N2	Real stock ownership of board members
N3	Change in stock ownership of board members
N4	Dominance of insider control
N5	Frequent CPA change

In this study, the sample covers 26 publicly traded firms encountering financial crises during the period from 2003 to 2007 in Taiwan while their non-distressed counterparts 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 208 (52*4=208) observations from the 4-year annual reports of the sampled firms, but only 196 observations are complete and available for use in the experiments. The 196 observations training data and A as testing data; and the third experiment uses A+C as training data and B as testing data. Training data are used for training and validation to select optimal parameters for the SVM and to prevent the over-fitting problem commonly found in the neural network method.

Besides, type I and type II errors were analyzed among these experiments. Type I error was defined as the probability that a firm predicted not to fail will in fact fail, while the Type II error was defined as the probability that a firm predicted to fail will not in fact fail.

Summary of profile analysis by features are 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 have 33% of the data used as a validation set.

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

Firm types	Distressed firm		Non-distressed firm	
Features	Mean	Std. Dev.	Mean	Std. Dev.
F1	13.67	17.59	18.78	13.12
F2	55.98	21.38	35.40	12.88
F3	0.00	0.35	0.00	0.17
N4	4.34	4.15	3.59	4.08
N5	0.00	0.32	0.00	0.17
N1	14.42	29.35	0.00	19.29
N3	-3.37	11.01	-1.53	5.84
N2	97.94	61.04	93.69	18.74

EXPERIMENT RESULTS AND DISCUSSION

PERFORMANCE COMPARISON

For performance comparison, we create three different prediction models: Model 1 based exclusively on our selected financial features; Model 2 based on attributes related to corporate governance; and Model 3, the new model we propose that combine both financial and non-financial features. Different types of errors result in different penalty costs. A comparative analysis of the errors reported in this study over a four-year period is presented. We discuss the three different models as follows.

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 (Long-term investment/total assets) and F2 (Total liabilities/total assets) emerge to be the more accurate of all the three financial predictors covered in Model 1. The predictive accuracy of Model 1 was then evaluated. As summarized in Table 4, the results of regression analysis suggest an average of 89.06% for the best prediction results. The model correctly classifies 88.24% of the non-distressed firms and 90.00% of the distressed firms. Predictions of distressed firms not to fail (Type I error) were greater than predictions of non-distressed firms to fail (Type II error). In Model 1, the average Type I and Type II errors are 11.76% and 10%, respectively.

Table 3: The Best predictive accuracy of Model 1

Model 1: Financial Features		Prediction D	Prediction N	Accuracy (%)
Observations	D	30	4 (Type I)	88.24%
	N	3 (Type II)	27	90.00%
Average				89.06%
D: distressed firm; N: non-distressed firm by using F1, F2 features				

Model 2 examines non-financial features to predict distressed firms with SVM. According to the results of the best experiment on model 2, N1 (Pledged shares of board members), N2 (Real stock ownership of board members), N3 (Change in stock ownership of board members), and N4 (Dominance of insider control) were found to be the more accurate of all the non-financial predictors covered in Model 2. Predictive accuracy of Model 2 is then evaluated, and the results indicate that the model accurately classifies 41.12% of the non-distressed firms and 93.33% of the distressed firms. The average of the best prediction results read 65.63%.

Table 4. The Best predictive accuracy of Model 2

Model 2: Non-Financial Features		Prediction D	Prediction N	Accuracy (%)
Observations	D	14	20 (Type I)	41.12%
	N	2 (Type II)	28	93.33%
Average				65.63%
D:distressed firm; N:non-distressed firm by using N1, N2, N3, N4 features				

Model 3, our proposal Model, combines both financial and non-financial features. Predictions of distressed firms not to fail (Type I error) were greater than predictions of non-distressed firms to fail (Type II error). In Model 3, the average Type I and Type II errors are 8.82% and 3.33%, respectively.

Based on the best experiment on model 3, F2 (Total liabilities/total assets), N2 (Real stock ownership of board members), N3 (Change in stock ownership of board members), and N4 (Dominance of insider control) emerge to be the more accurate of all the financial and non-financial predictors covered in Model 2. Predictive accuracy of Model 3 is also evaluated, and the results revealed that the model accurately classifies 91.11% of the non-distressed firms and 96.67% of the distressed firms. The average of the best prediction results read 93.75%, which is significantly superior to both Model 1 and Model 2. Empirical results show that our SVM model examining both financial and non-financial features can serve as a promising alternative for existing financial distress prediction models.

Table 5. The Best predictive accuracy of Model 3

New Model: Mixed (Financial + Non-Financial Features)		Prediction D	Prediction N	Accuracy (%)
Observations	D	31	3 (Type I)	91.11%
	N	1 (Type II)	29	96.67%
Average				93.75%
D:distressed firm; N:non-distressed firm by using F2, N2, N3, N4 features				

As Table 6 shows, the average predictive accuracy of all three models was between 78.34% to 83.95% in predicting companies in the TSE market that failed. Model 3 of our proposal Model is able to predict bankruptcy (83.95% accuracy). Non-financial Model (Model 2) exhibited the lowest predictive accuracy of all the models. The combination of financial features and non-financial features outperformed Model I and Model 2.

Table 6: Predictive accuracies of models in holdout sample (the TSE market)

Evaluation criterion	Financial (Model 1)	Non-financial (Model 2)	Combination (Model 3)
Type I error	0.2656	0.5965	0.1989
Type II error	0.1713	0.1439	0.1194
Brier Score (BS)	0.2166	0.3732	0.1605
Average predictive accuracy	0.7834	0.6268	0.8395
Best feature selection	[F2]	[N1],[N3]	[F2],[N4],[N2],[N3]

*the experiment using cross-validation

In general, financial status of listed companies in Taiwan can be better predicted using the proposed SVM model since the average predictive accuracy of the failing company model is as high as 83.95% by using both financial and non-financial features. The proposed model (Model 3) outperformed other bankruptcy models. In practice, the cost of misclassifying a failed firm is likely to be much greater than that of misclassifying a non-failed firm. Type I is the probability of misclassifying a failed while Type II error is the probability of misclassifying a non-failed firms. In Model 1 (financial feature only) predictions of failed firms not to fail (Type I error) were greater than predictions of non-failed companies to fail (Type II error). The average Type I and II errors are 26.56% and 17.13%, respectively. In Model 2 (non-financial feature only), the average Type I and II errors are 59.65% and 14.39%, respectively. In the proposed SVM model (Model 3) betrays a classification error ratio of 19.89% (Type I errors). This may be partly due to the limitation of all classification models when they are applied to the overlapping financial data of two groups, and partly because of the inclusion of "full-value delivery" as one of the criterion for business crises. Whether to impose the special treatment of full-value delivery on listed companies can be problematic since, on the one hand, a genuinely financially distressed company may escape the special treatment through profit manipulation while, on the other hand, a financially solid company may receive the special treatment through negligence or accident. In spite of the above mentioned concerns, full-value delivery event remains a viable cut-off event for listed companies in Taiwan if it can be easily observed and carefully evaluated.

We also adopted Brier Score (BS) [4] to compare different prediction accuracy. The Brier Score (BS) is a measure of prediction accuracy that is a well-known in meteorology and medical science [4]. It is calculated as [$BS = \frac{1}{n}\sum_{i}^{n}(\theta_i - p_i)^2$] where θi is a binary indicator for the actual realization of the default variable (1 if default, 0 if no default) and pi, 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 6, the combination of financial and nonfinancial features achieves a lower average Brier Score (BS) of 16.05% after taking into account of all experiment results. F2 for its omnipresence in the combinations apparently plays the most important role. Of the non-financial features, the importance of N3 as a leading indicator is testified by the fact that it is the most frequently detected non-financial features in the combinations. The data summarized in Table 6 are further consulted to identify an optimal feature set defined as the one capable of reaching the highest possible prediction accuracy rate with the lowest possible number of features. And the result, as indicated by the table 6, goes to the feature set 1 of F2 (Total liabilities/total assets), N4 (Dominance of insider control), N3 (Change in stock ownership of board members), and N2 (Real stock ownership of board members).

In summary, a "mixed" model encompassing both financial and non-financial features leads to a more accurate prediction of corporate financial distress than a model based exclusively on either non-financial or financial feathers.

CONCLUSION

This paper constitutes an attempt to explore both the role of financial and non-financial feathers related to corporate governance in business crisis prediction. The SVM model developed by our study to assess corporate bankruptcy risk demonstrates significantly improved accuracy over existing business distressed classification models. The test results by our proposed model report. An overall predictive accuracy rate of 83.95% is superior to both the financial and non-financial feature only for business crisis prediction. Moreover, since the extensive presence of concentrated ownership in public listed companies in Taiwan has rendered corporate governance a crucial concern in business crisis prediction, we analyze non-financial features related to corporate governance, notably pledged shares of board members, real stock ownership of board members, and change in stock ownership of board members, via the SVM method. Inclusion of these non-financial features appears to help 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. Additionally, the machine-learning models are more accurate in predicting financial distress than other multivariate statistical models.

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 prediction 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 role of ownership structure and corporate governance play in financial distress 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 since the inclusion of non-financial features in business crisis prediction has been proved to enhance prediction performance in our study.

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.” In 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 78*.pp1-3, (1950).
5. Burbidge, *Nancy M.*, Ph.D., ‘‘Organ allocation and patient responsibility: Re-examining the concept of responsibility in light of the thought of Emmanuel Levinas’’, *University of Ottawa (Canada)***,** 337 pages; AAT NQ66129, (2001).
6. Chang, C. C. and Lin, C. J., “LIBSVM: A library for support vector machines,” (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*, 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***.** Cambridge: Dec. Vol. 57, Iss. 6; p. 2741, (2002).
10. Claessens, S., Djankov, S. and Xu, L. C., ‘‘Corporate performance in the East Asian financial crisis,’’ *The World Bank Research Observer.* Cary: Feb. Vol. 15, Iss. 1, p. 23, (2000).
11. Eichengreen, B.‘‘Is greater private sector burden sharing impossible?,’’ *Finance & Development***,** Washington: Sep. Vol. 36, Iss. 3; p. 16, (1999).
12. Gilson and Stuart C, ‘‘Management Turnover and Financial Distress,’’ *Journal of Financial* Economics, Amsterdam: Dec. Vol. 25, Iss. 2, p. 241, (1989).
13. Günther, T., Grüning, M., 2000. Einsatz von Insolvenzprognoseverfahren bei der Kreditwürdigkeitsprüfung im Firmenkundenbereich. Die Betriebswirtschaft 60, 39–59
14. 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*, 37 ,543-558 , (2004)
15. http://www.kpmg.com/xbrl/XBRL_Home.asp
16. Lee, K. C., Han, I., and Kwon, Y., “Hybrid neural network models for bankruptcy predictions, *Decision Support Systems*, 18 (1),

63-72, (1996).

17. Lee, T. S., Woidtke, T., and Yeh, Y. H., "Corporate governance and financial distress: evidence from Taiwan," *Corporate Govrnance*, Vol. 12, pp. 378-388, (2004).
18. Lien, Y. Piesse, J., Strange, R. and Filatotchev, I., "The role of corporate governance in FDI decisions: Evidence from Taiwan," *International Business Review*, pp 739-763, 14 (2005).
19. Mao, H., Carson, J.M., Ostaszewski, K. M. and Shoucheng, L,"Pricing Life Insurance: Combining Economic, Financial, and Actuarial Approaches", *Journal of Insurance Issues*. Northridge: Fall. Vol. 27, Iss. 2; p. 134 (2004).
20. 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)
21. Ohlson, J.A., "Financial ratios and the probabilistic prediction of bankruptcy," *Journal of Accounting Research*, Vol. 18, pp. 109-31, Spring (1980).
22. 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).
23. Sharmaand, S. and Mahajan, V., "Early warning indicators of business failure," *Journal of Marketing*, fall, pp. 80-9.
24. 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).
25. SPSS 13.0, url://www.spss.com
26. Taffler, R. and Tisshaw, H., "Going, going, gone-four factors which predict," *Accountancy*, Vol. 88, pp. 50-4, March (1977).
27. Tam, K. Y., & Kiang, M.Y., Managerial applications of neural networks: the case of bank failure predictions. Management Science, 38 (7), 926-947, (1992).
28. 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*, 32, 397-408, (2007).
29. Yeh, Y. H., and Woidtke, T., "Commitment or entrenchment? Controlling shareholders and board composition," *Journal of Banking & Finance*, 29, 1857-1885, (2005).