Available online at www.sciencedirect.com

ScienceDirect

Expert Systems with Applications 32 (2007) 397–408

Expert Systems with Applications

www.elsevier.com/locate/eswa

# A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy

Chih-Hung Wu [a], Gwo-Hshiung Tzeng [b,c,*], Yeong-Jia Goo [d], Wen-Chang Fang [d]

[a] Department of Business Administration, Takming College, No. 56, Sec.1, Huanshan Rd., Neihu District, Taipei City 114, Taiwan
[b] Department of Business Administration, Kainan University, No. 1, Kainan Rd., Luchn, Taoyuan 338, Taiwan
[c] Institute of Management of Technology, National Chiao Tung University, 100 Ta-Hsueh Rd., Hsinchu 300, Taiwan
[d] Department of Business Administration, National Taipei University, Taipei, Taiwan

**Abstract**

Two parameters, $C$ and $\sigma$, must be carefully predetermined in establishing an efficient support vector machine (SVM) model. Therefore, the purpose of this study is to develop a genetic-based SVM (GA-SVM) model that can automatically determine the optimal parameters, $C$ and $\sigma$, of SVM with the highest predictive accuracy and generalization ability simultaneously. This paper pioneered on employing a real-valued genetic algorithm (GA) to optimize the parameters of SVM for predicting bankruptcy. Additionally, the proposed GA-SVM model was tested on the prediction of financial crisis in Taiwan to compare the accuracy of the proposed GA-SVM model with that of other models in multivariate statistics (DA, logit, and probit) and artificial intelligence (NN and SVM). Experimental results show that the GA-SVM model performs the best predictive accuracy, implying that integrating the RGA with traditional SVM model is very successful.


*Keywords:* Support vector machine (SVM); Real-valued; Genetic algorithm (GM); Financial distress; Prediction; Bootstrap simulation

## 1. Introduction

Predicting corporate failure has been an important research topic in accounting and finance for the last three decades (Lee, Han, & Kwon, 1996; Salcedo-Sanz, Fernandez-Villacanas, Segovia-Vargas, & Bousono-Calzon, 2005). The financial crisis in East Asia provoked particularly extensive studies of the financial distress of institutions with various financial and ownership structures that arose across countries with very diverse institutional setups in East Asia in 1997 and 1998 (Claessens, Djankov, & Klapper, 2003). Classical studies on ratio analysis and the classification of bankruptcy was performed by Beaver's dichotomous classification test in 1967. Altman (1968) proposed the $Z$-score model, which applied multivariate discriminant analysis (MDA) and employed financial ratios as input variables to predict financial distress. Subsequent studies have developed more precise model to predict bankruptcy. Deakin (1972) revised Altman and Beaver's studies, using a quadratic function to construct a more precise classification model of financial distress and thus increase the accuracy for predicting financial distress. After that, logit regression (Ohlson, 1980; Platt & Platt, 1990; Tseng & Lin, 2005; Zavgren, 1985) or probit regression (Zmijewski, 1984) have widely adopted in subsequent work. Nevertheless, empirical results have shown that most of financial ratios violate the assumptions of the multivariate statistical model used in these previous studies. In recently studies, several revised financial distress models such as the revised the $Z$ score and ZETA models and the hybrid system (Lee et al., 1996; Tam & Kiang, 1992) have been demonstrated the results of highly adaptable and outperformed in predicting bankruptcy. In addition, statistical learning

* Corresponding author. Fax: +886 3 3412430.
*E-mail addresses:* ericwu@mail.takming.edu.tw (C.-H. Wu), ghtzeng@cc.nctu.edu.tw, ghtzeng@mail.knu.edu.tw (G.-H. Tzeng), goo@ntpu.edu.tw (Y.-J. Goo), fang@ntpu.edu.tw (W.-C. Fang).

0957-4174/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2005.12.008

approaches such as neural networks or support vector machines (SVM) has been successfully applied to this kind of problems (Min & Lee, 2005; Salcedo-Sanz et al., 2005; Tam, 1991; Tam & Kiang, 1992; Wu, 2004).

Previous application of neural networks in finance and accounting, notably in bankruptcy prediction, are limited to back-propagation neural networks (Yang, Platt, & Platt, 1999). Recently, new algorithms in machine learning, Support vector machines (SVMs), were developed by Boster, Guyon, and Vapnik (1992) to provide better solutions to decision boundary than could be obtained using the traditional neural network. Since the new model was proposed (Boster et al., 1992; Cortes & Vapnik, 1995), SVM has been successfully applied to numerous applications, including the handwriting recognition, particle identification (e.g., muons), digital images identification (e.g., face identification), text categorization, bioinformatics (e.g., gene expression), function approximation and regression, time series forecasting (Cao, 2003; Kim, 2003; Mukherjee, Osuna, & Girosi, 1997; Müller, Smola, Rätsch, & Schölkopf, 1999; Tay & Cao, 2001, 2002), chaotic system (Mattera & Haykin, 1999) and bankruptcy prediction (Min & Lee, 2005; Salcedo-Sanz et al., 2005). This study examines the possibility of enhancing the accuracy of predicting bankruptcy by adopting the SVM model.

Min and Lee (2005) stated that the optimal parameter search on SVM plays a crucial role to build a bankruptcy prediction model with high prediction accuracy and stability. To make an efficient SVM model, two extra parameters: $C$ and $\sigma^2$ (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, $\sigma^2$, is the bandwidth of the radial basis function (RBF) kernel. Consequently, the purpose of this study is to propose a model that can determine the optimal parameters ($C$ and $\sigma^2$) of SVMs to yield the highest predictive accuracy and generalization ability for predicting bankruptcy. The model was tested on the prediction of financial crisis of Taiwan to compare its accuracy with that of other models that based on multivariate statistics and AI approaches.

The remainder of this paper is organized as follows. The basic ideas of methods for bankruptcy prediction is reviewed and discussed in Section 2. Research design for modeling genetic based SVM is proposed in Section 3 to describe its ideas and procedures. An example of empirical analysis for predicting bankruptcy is used to demonstrate the proposed method in Section 4. Discussions are presented in Section 5 and conclusions are in the last section.

## 2. Basic ideas of methods for bankruptcy prediction

In this section, the basic ideas of methods for bankruptcy prediction from the perspective of the non-linear SVM are provides. Then, the real-valued genetic algorithm is briefly introduced. Parameters optimization approaches are discussed in the following section. Finally, statistical approaches for predicting bankruptcy are overviewed in the final section.

### 2.1. The non-linear support vector machine

The basic idea in designing a non-linear SVM model is to map the input vector $\mathbf{x} \in \Re^n$ into vectors $\mathbf{z}$ of a higher-dimensional feature space $\boldsymbol{F}$ ($\mathbf{z} = \boldsymbol{\varphi}(\mathbf{x})$, where $\boldsymbol{\varphi}$ denotes the mapping $\Re^n \to \Re^f$), and to solve a linear classification problem in this feature space

$$\mathbf{x} \in \Re^n \to \mathbf{z}(\mathbf{x}) = [a_1\boldsymbol{\varphi}_1(\mathbf{x}), a_2\boldsymbol{\varphi}_2(\mathbf{x}), \ldots, a_n\boldsymbol{\varphi}_n(\mathbf{x})]^{\mathrm{T}} \in \Re^f. \tag{1}$$

Namely, the basic idea in non-linear SVM is to map the data $\mathbf{x}$ into a high-dimensional feature space via a mapping function $\boldsymbol{\varphi}(\mathbf{x})$ (also called kernel function), which is selected by the user in advance. By replacing the inner product for non-linear pattern problem, the kernel function can perform a non-linear mapping to a high-dimensional feature space (Vapnik, 1995). Kernel functions perform the non-linear mapping between input space and a feature space.

The approximating feature map for the Mercer kernel is $K(\mathbf{x},\mathbf{y}) = \boldsymbol{\varphi}(\mathbf{x})^{\mathrm{T}}\boldsymbol{\varphi}(\mathbf{y})$, which performs the non-linear mapping. Currently, popular kernel functions in machine learning theories are as follows (Campbell, 2002; Kecman, 2001).

$$\text{Gaussian (RBF) kernel}: \; K(\mathbf{x}_i,\mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i-\mathbf{x}_j\|^2}{2\sigma^2}\right); \tag{2}$$

$$\text{Polynomial kernel:} \; K(\mathbf{x}_i,\mathbf{x}_j) = (1+\mathbf{x}_i^{\mathrm{T}}\mathbf{x}_j)^d; \tag{3}$$

$$\text{Linear kernel:} \; K(\mathbf{x}_i,\mathbf{x}_j) = \mathbf{x}_i^{\mathrm{T}}\mathbf{x}_j; \tag{4}$$

$$\text{Multilayer perceptron:} \; K(\mathbf{x}_i,\mathbf{x}_j) = \tanh[(\mathbf{x}^{\mathrm{T}}\mathbf{x}_i)+b]. \tag{5}$$

In Eq. (2), $\sigma^2$ denotes the variance of the Gaussian kernel. A certain value of $b$ is used only in the multilayer perceptron.

The learning algorithm for a non-linear classifier SVM follows the design of an optimal separating hyperplane in a feature space. The procedure the same as associated with hard and soft margin classifier SVMs in $x$-space. Accordingly, the dual Lagrangian in $z$-space is

$$L_{\mathrm{d}}(\boldsymbol{\alpha}) = \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} y_i y_j \alpha_i \alpha_j \mathbf{z}_i^{\mathrm{T}}\mathbf{z}_j, \tag{6}$$

and using the chosen kernels, the Lagrangian is maximized as follows.

$$\text{Maximize:} \quad L_{\mathrm{d}}(\boldsymbol{\alpha}) = \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} y_i y_j \alpha_i \alpha_j K(\mathbf{x}_i\mathbf{x}_j) \tag{7}$$

$$\text{Subject to} \quad \alpha_i \geqslant 0, \quad i = 1,\ldots,l, \tag{8}$$

$$\sum_{i=1}^{l}\alpha_i y_i = 0. \tag{9}$$

Note the constraints must be revised for using in a non-linear soft margin classifier SVM. The only difference these constraints and those of the separable non-linear classifier are in the upper bound $C$ on the Lagrange multipliers $\alpha_i$. Consequently, the constraints of the optimization problem become

$$\text{Subject to} \quad C \geqslant \alpha_i \geqslant 0, \quad i = 1, \ldots, l, \tag{10}$$

$$\sum_{i=1}^{l} \alpha_i y_i = 0. \tag{11}$$

In this way, the influence of the training data point will be limited and remained on the wrong side of a separating non-linear hypersurface. The decision hypersurface $d(\mathbf{x})$ and the indicator function, which were determined by the non-linear SVM classifier, are as follows:

$$d(\mathbf{x}) = \sum_{i=1}^{l} y_i \alpha_i K(\mathbf{x}_i, \mathbf{x}_j) + b, \tag{12}$$

$$i_F(\mathbf{x}) = \text{sign}(d(\mathbf{x})) = \text{sign}\left(\sum_{i=1}^{l} y_i \alpha_i K(\mathbf{x}_i, \mathbf{x}_j) + b\right). \tag{13}$$

Depending upon the chosen kernel, the bias term $b$ may implicitly be a part of the kernel function. For example, the bias term $b$ is not required when Gaussian RBFs are used as kernels. When the bias term $b$ is included within other kernel functions, the non-linear SVM classifier is as follows:

$$i_F(\mathbf{x}) = \text{sign}(d(\mathbf{x})) = \text{sign}\left(\sum_{i=1}^{l} y_i \alpha_i K(\mathbf{x}_i, \mathbf{x}_j) + b\right) = \text{sign}\left(\sum_{s=1}^{\text{number of SVs}} y_s \alpha_s K(\mathbf{x}, \mathbf{x}_s)\right). \tag{14}$$

### 2.2. Real-valued genetic algorithm (RGA)

Recently, genetic algorithms (GAs) have been widely and successfully applied to various optimization problems (Fogel, 1994; Goldberg, 1989; Grefenstette, 1986). GAs are well suited to the concurrent manipulating of models with varying resolutions and structures since they can search non-linear solution spaces without requiring gradient information or a priori knowledge about model characteristics (McCall & Petrovski, 1999). The problem existing in the binary coding lies in the fact that a long string always occupies the computer memory even though only a few bits are actually involved in the crossover and mutation operations. This is particularly the case when a lot of parameters are needed to be adjusted in the same problem and a higher precision is required for the final result. To overcome the inefficient occupation of the computer memory, the underlying real-valued crossover and mutation algorithms are employed (Huang & Huang, 1997). In contrast to the binary genetic algorithm (BGA), 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 (Haupt & Haupt, 1998). Namely, RGA is more straightforward, faster and more efficient than BGA. Since this study is concerned with finding optimal values of SVM parameters whose precise values are unknown, the aforementioned properties of RGA are highly advantageous.

### 2.3. Parameter optimization

To design an effective SVM model, values of parameters in SVM have to be chosen carefully in advance (Duan, Keerthi, & Poo, 2003; Lin, 2001; Min & Lee, 2005). These parameters include the following: (1) regularization parameter $C$, which determines the tradeoff cost between minimizing the training error and minimizing the complexity of the model; (2) parameter sigma ($\sigma$ or $d$) of the kernel function which defines the non-linear mapping from the input space to some high-dimensional feature space. This investigation only considers only the Gaussian kernel, the variance of whose function is sigma squared $\sigma^2$; (3) a kernel function used in SVM, which constructs a non-linear decision hypersurface in an input space.

Cristianini, Shawe-Taylor, and Campell (1998) proposed the Kernel-Adatron Algorithm which can automatically select models without testing on a validation data. Unfortunately, this algorithm is ineffective if the data have a flat ellipsoid distribution (Campbell, 2002). Therefore, one possible way to solve the problem is to consider the distribution of the data. Interestingly, various specific functions in SVM, after the learning stage, can create the decision hypersurfaces of the same type (Kecman, 2001).

To solve the problem, Lin (2001) provided a systematic method for selecting SVM parameters. His systematic design for selecting parameters of support vector regression was adopted the concept of the sampling theory into Gaussian Filter. Min and Lee (2005) proposed a grid-search technique using 5-fold crossvalidation to find out the optimal parameters values of kernel function of SVM.

In contrast to abovementioned methods of parameter optimization on SVM, this study develops a new method, named GA-SVM, for optimizing the two SVM parameters ($C$ and $\sigma^2$) simultaneously. The first parameter, $C$, determines the trade-off between the fitting error minimization and model complexity. The second parameter, $\sigma^2$, is the bandwidth of the radial basis function (RBF) kernel.

### 2.4. Overview of statistical approaches for predicting bankruptcy

The corporate distress literature includes several diverse methodologies for discriminating between failed and non-failed firms, following Beaver's univariate comparison of financial ratios in 1966. Extensive studies in this area have applied statistical and AI approaches over the last three decades. The well-known multivariate models used in this area include multiple discriminate analysis (MDA) (Altman, 1968; Altman, Haldeman, & Narayanan, 1977), logit analysis (Ohlson, 1980; Platt & Platt, 1990; Tseng &

Lin, 2005; Zavgren, 1985), and probit analysis (Zmijewski, 1984). Most recently, AI approaches, such as neural network approaches (Lee et al., 1996; Lee, Booth, & Alam, 2005; Odom & Sharda, 1990; Tam, 1991; Yang et al., 1999) or SVM (Min & Lee, 2005; Salcedo-Sanz et al., 2005) have shown promise as classification tools.

## 3. Designing a genetic-based SVM model for predicting bankruptcy

In this section, we describe the design of our proposed model, a genetic-based SVM model, for predicting bankruptcy. The approach of combining real-valued genetic algorithm with SVM is introduced in the first section. Research data and description of samples are described in the next section. Modeling and the parameter settings of BPN and SVM are presented in Section 3. Chromosome representations, the design of fitness function and genetic operators in this study are discussed in the final sections.

### *3.1. Our proposed approach*

In the proposed GA-SVM model, the SVM parameters are dynamically optimized by implementing the RGA evolutionary process and the SVM model then performs the prediction task using these optimal values. Namely, the RGA tries to search the optimal values to enable SVM to fit various datasets. The process of GA-SVM was illustrated in Fig. 1. The optimal values of SVM's parameters are searching by GAs with a randomly generated initial populations consisting of chromosomes. The values of the two parameters, $C$ and $\sigma^2$, are directly coded in the chromosomes with real-valued data. The proposed model can implement either the roulette-wheel method or the tournament method for selecting chromosomes. Adewuya's crossover method and boundary mutation method were used to modify the chromosome. The single best chromosome in each generation is survives to the succeeding generation. The proposed model was developed and implemented in the MATLAB v6.5 environment. The major tool for training and validating the SVM were those developed by Pelckmans et al. (2002). The proposed model is able to handle huge data sets and easily be combined with the real-valued genetic algorithm in the MATLAB environment.

Predicting bankruptcy (or financial distress) has been a major research issue in accounting and finance over in the last three decades (Lee et al., 1996). Therefore, the genetic-based SVM (GA-SVM) model was applied to the problem of financial distress in Taiwan to verify its accuracy and generalization ability, must be shown to be more accurate than the traditional multivariate statistical models and neural network technique.

### *3.2. Research data*

Financial-statement data of the failed and non-failed firms were obtained from the database of the Taiwan Economic Journal (TEJ), covering in cases of 3 years prior to failure and 1 year after failing. "Failure" is defined as the inability of a firm to pay its financial obligations as they mature. A firm is specifically said to have failed when any of the following events have occurred: bankruptcy, default on bonds, the overdrawing of a bank account, or non-payment of a preferred stock dividend (Beaver, 1966). This study defined the firms in financial distress as those whose listed securities have been classified as the category of alter-trading-method.[1] When any of the events exists in aforementioned events occurs in the operating rules, this Corporation may place the listed securities under the category of altered-trading-method.

According to the definition of Beaver (1966), the "first year before failure" is defined as that year included in the most recent financial statement prior to the year in which the firm is reported to have failed. The data sample consists of firms in Taiwan that failed in the period from 1998 to 2002. The failed firms were selected from the lists of bankrupt companies by the Taiwan stock exchange (TSE) and the database of TEJ. A failed firm was paired with a non-failed firm by (1) industry, (2) products, (3) capitalization, and (4) values of assets. Table 1 presents the description of samples. Failed companies were paired with non-failed firms in a similar industry, dealing in similar products, with similar capitalization, and with similar values of assets.

The size of matched sample was 88 firms, including 22 failed firms and 66 non-failed firms. In the simulated sample, the total sample size was 44 companies, including 22 failed firms and 22 non-failed firms. The holdout sample comprises of all corporations listed on the TSE and OTC market from 2001 to 2002. The sample size for 2001 was 538 firms, including 373 firms on the TSE and OTC market in 2001. The sample size for 2002 was 534 firms, including 356 firms on the TSE and OTC market.

Table 1 also presents further details about the matched sample in this study. The matched sample was paired according to industry, primary product, capitalization and values of assets. A lower matching ratio (e.g., 1:1 or 1:2) corresponds to higher bias in the selection in choice-based, which leads to oversampling (Platt & Platt, 2002; Zmijewski, 1984). Using the matching rule that has been proposed by Su (2000), this study adopted one financially distressed firm was paired to three non-failed firms (1:3 ratio) to avoid the problem of oversampling and bias in the choice-based sample (Platt & Platt, 2002).

### *3.3. Modeling*

The 19 financial variables are those which have been found or actually used in previous research to be significant in predicting bankruptcy. These ratios can be grouped into

[1] This method is according to Articles 49, 50, and 50-1 of the Operating Rules of the Taiwan Stock Exchange Corporation.

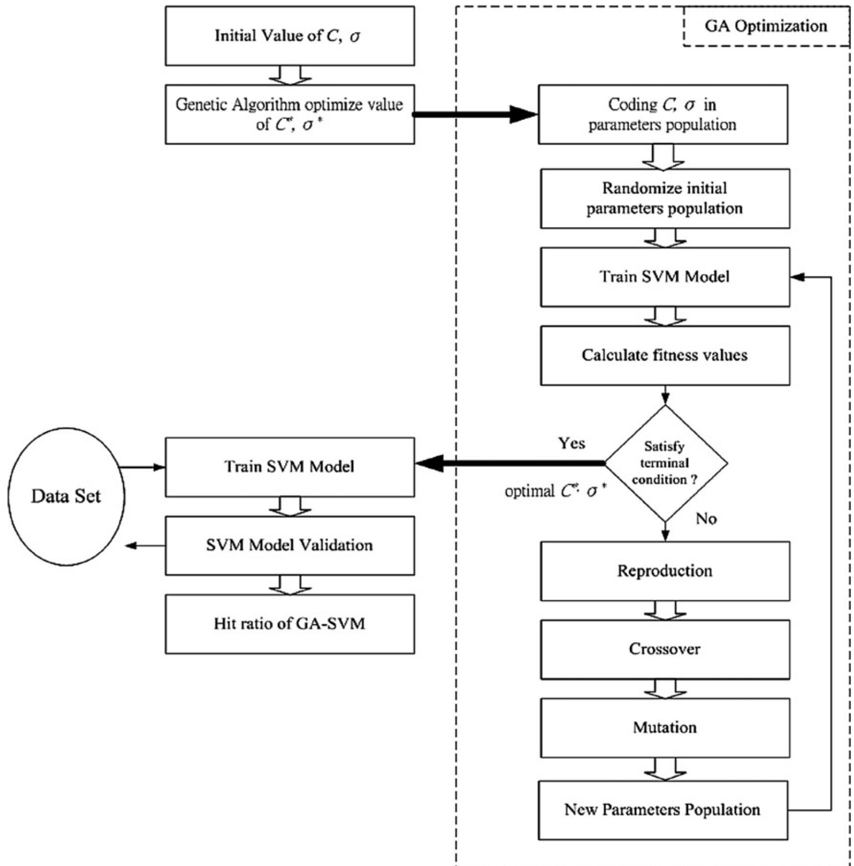


Fig. 1. The process of this approach.

Table 1
Description of samples

| Samples | Paired rate (F vs. NF) | Number of observations F | NF | Total | Time period |
|---|---|---|---|---|---|
| Matched sample | 1:3 | 22 | 66 | 88 (training and forecasting) | 1998–2000 |
| Simulated sample (random sampling by bootstrap) | 1:1 | 15 | 15 | 30 (training) | 1998–2000 |
| | | 7 | 7 | 14 (forecasting) | |
| Holdout sample | None | 7 | 531 | 538 (forecasting for TSE in 2001) | 2001–2002 |
| | | 14 | 520 | 534 (forecasting for TSE in 2002) | |
| | | 5 | 368 | 373 (forecasting for OTC in 2001) | |
| | | 5 | 351 | 356 (forecasting for OTC in 2002) | |

*Note:* F = failed firms; NF = non-failed firms.

four categories, including liquidity, profitability, asset management, and financial structure. Initially, financial variables are selected for used in the bankruptcy prediction model. A list of tested financial ratios is summarized in Table 2.

### 3.3.1. Neural network

The feed-forward back-propagation neural network (BPN) applied to the experimental sample includes 19 input neurons in the input layer, seven neurons in the hidden layer, and one in the output layer. This study constructed a three-layer network and employed the "TRAINLM algorithm", "LEARNGDM", and "MSEREG" as the training function, the adaptive learning function, and the performance function, respectively. The transfer function was set to the "TANSIG function" and the "PURELIN function" for hidden layer and output layer, respectively. The number of epochs was set to 300 and the learning rate was set to 0.05 in each epoch. Table 3 presents the parameter settings.

### 3.3.2. The SVM model

When data sets are noisy and exhibit a large overlap between data classes, error variables $\varepsilon_i > 0$ are introduced

Table 2
List of tested financial ratios

| Section | Financial ratios |
|---|---|
| Liquidity | Current ratio<br>Quick ratio<br>Cash flow ratio |
| Profitability | Net income to sales<br>Gross profit to sales<br>Net income to total assets<br>Net income to stockholder's equity<br>Operating income to sales<br>Earning per share (EPS)<br>Growth ratio of sales |
| Asset management | Total asset turnover<br>Fixed assets turnover<br>Inventory turnover<br>Receivables turnover |
| Financial structure | Debt ratio<br>Long-term liabilities to fixed assets<br>Degree of financial leverage (DFL)<br>Liabilities to stockholder's equity<br>Interest coverage ratio |

Table 3
Parameter settings used in BPN

| Parameters of NN | Values |
|---|---|
| Network type | FEED-FORWARD BACK-PROPAGATION |
| Training function | TRAINLM |
| Adaptive learning function | LEARNGDM |
| Performance function | MSEREG |
| Number of layers | 3 |
| Neurons in hidden layer | 7 |
| Transfer function of hidden layer | TANSIG |
| Transfer function of output layer | PURELIN |
| Epochs | 300 |
| Learning rate | 0.05 |

to allow the output of the outlier to be locally corrected, constraining the range of the Lagrange multiplier $\alpha_i$ from 0 to $C$. $C$ is a constant penalty function designed to prevent outliers from affecting the optimal hyperplane. Hence, the non-linear objective function is

$$\text{Maximize:} \quad W(\boldsymbol{\alpha}) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j (K(\mathbf{x}_i, \mathbf{x}_j)) \tag{15}$$

$$\text{Subject to} \quad 0 \leqslant \alpha_i \leqslant C, \quad i = 1, \ldots, l, \tag{16}$$

$$\sum_{i=1}^{l} \alpha_i y_i = 0. \tag{17}$$

The optimal weight $\mathbf{w}^*$ and bias are determined by solving the quadratic programming problem.

$$\mathbf{w}^* = \sum_{i=1}^{l} \alpha_i^* y_i \mathbf{x}_i, \tag{18}$$

$$b^* = y_i - \mathbf{w}^{*\mathrm{T}} \mathbf{x}_i \tag{19}$$

The optimal decision function is as follows:

$$f(\mathbf{x}) = \operatorname{sign}\left( \sum_{i=1}^{l} y_i \alpha_i^* K(\mathbf{x}, \mathbf{x}_i) + b^* \right). \tag{20}$$

In machine learning theories, popular kernel functions, such as the Gaussian kernel function, have been found to provide good generalization capabilities (Campbell, 2002; Kecman, 2001). Accordingly, the Gaussian kernel function is employed as the kernel function in this work. The Gaussian kernel function is given by

$$\text{Gaussian (RBF) kernel:} \quad K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left( -\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2} \right). \tag{21}$$

A kernel function $K(\mathbf{x}, \mathbf{z})$, satisfying Mercer's condition, performs a high dimensional mapping $\boldsymbol{\varphi} : \Re^N \mapsto F$ and be used as a substitute for $(\boldsymbol{\varphi}(\mathbf{x}) \cdot \boldsymbol{\varphi}(\mathbf{z}))$ which replaces $(\mathbf{x} \cdot \mathbf{z})$ (Vapnik, 1995). Consequently, the optimal hyperplane classification function is obtained by the SVM model to fit an optimal hyperplane between two classes in a training data set.

### 3.4. Chromosome representations

Unlike the traditional BGA, the RGA used to solve optimization problems, directly codes all of the corresponding parameters or variables in a chromosome. Hence, the representation of the chromosome is straightforward in the RGA. The two parameters, $C$ and $\sigma$, of SVM were directly coded to form the chromosome in the proposed method. The chromosome $X$ is represented as $X = \{p_1, p_2\}$, where $p_1$ and $p_2$ denote the regularization parameter $C$ and sigma $\sigma$ (the parameter of the kernel function), respectively.

### 3.5. The fitness function

A fitness function, assessing the performance of each chromosome, must be designed before starts to search optimal values of SVM parameters. Several measurement indicators have been developed and applied to evaluate the predictive accuracy of models; they include the hit ratio, MAPE, RMSE, and the maximum error. The hit ratio is used herein as the indicator of model performance to compare the results achieved by the proposed model with those obtained using other models (traditional SVM, discriminant analysis, logit analysis, probit regression, and NN). The hit ratio denotes the value of the fitness function in GA-SVM.

### 3.6. Genetic operators

The real-valued genetic algorithm uses selection, crossover, and mutation operators to generate the offspring of the existing population.

#### 3.6.1. Selection

The proposed GA-SVM model incorporates two well-known selection methods – the roulette wheel method and the tournament method. The tournament selection method is adopted here to decide whether a chromosome can survive to the next generation. The chromosomes that survive to the next generation are placed in a matting pool for crossover and mutation operations.

#### 3.6.2. Crossover

Once a pair of chromosomes has been selected for crossover, one or more randomly selected positions are assigned to the to-be-crossed chromosomes. The newly crossed chromosomes are then combined with the rest of the chromosomes to generate a new population. However, overloading problem frequently occurs when the RGA is used to optimize values. This study uses the method proposed by Adewuya (1996) to prevent overload of post-crossover when genetic algorithm with real-valued chromosomes are applied.

$$X_1^{\text{old}} = \{x_{11}, x_{12}, \ldots, x_{1n}\}, \quad X_2^{\text{old}} = \{x_{21}, x_{22}, \ldots, x_{2n}\} \tag{22}$$

move closer:

$$X_1^{\text{new}} = X_1^{\text{old}} + \sigma(X_1^{\text{old}} - X_2^{\text{old}}), \tag{23}$$

$$X_2^{\text{new}} = X_2^{\text{old}} - \sigma(X_1^{\text{old}} - X_2^{\text{old}}) \tag{24}$$

move away:

$$X_1^{\text{new}} = X_1^{\text{old}} + \sigma(X_2^{\text{old}} - X_1^{\text{old}}), \tag{25}$$

$$X_2^{\text{new}} = X_2^{\text{old}} - \sigma(X_2^{\text{old}} - X_1^{\text{old}}). \tag{26}$$

$X_1^{\text{old}}$ and $X_2^{\text{old}}$ represent the pair of populations before crossover operation; $X_1^{\text{new}}$and $X_2^{\text{new}}$ represent the pair of new populations after crossover operation. In addition, $\sigma$ is a random micro number that controls the variance of each crossover operations.

#### 3.6.3. Mutation

The mutation operation follows the crossover operation and determines whether a chromosome should be mutated in the next generation. In this study, uniform mutation method is applied and designed in the presented model. Consequently, researchers can select the method of mutation in GA-SVM best suited to their problems of interest.

Uniform mutation

$$X^{\text{old}} = \{x_1, x_2, \ldots, x_n\}, \tag{27}$$

$$x_k^{\text{new}} = \text{LB}_k + r^*(\text{UB}_k - \text{LB}_k), \tag{28}$$

$$X^{\text{new}} = \{x_1, x_2, \ldots, x_k^{\text{new}}, \ldots, x_n\}, \tag{29}$$

where $n$ denotes the number of parameters; $r$ represents a random number in the range $(0,1)$, and $k$ is the position of the mutation. LB and UB are the low and upper bounds on the parameters, respectively. $\text{LB}_k$ and $\text{UB}_k$ denote the low and upper bound at location $k$. $X^{\text{old}}$ represents the population before mutation operation; $X^{\text{new}}$ represents the new population following mutation operation.

## 4. Empirical analysis for predicting bankruptcy

Empirical analysis is divided into five sections: (1) data analysis, (2) normality testing, (3) predictive accuracy of matched samples, (4) predictive accuracy of holdout sample, and (5) predictive of simulated (bootstrap) sample.

### 4.1. Data analysis

Experiments were performed to examine three kinds of validation: (1) internal validation (matched sample), (2) external validation (holdout sample prediction) and (3) external validation (simulated sample prediction). Besides the accuracy of the predictions of bankruptcy, 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 (Blum, 1974). The SVM model is applied with fix values of parameters (Fig. 2).

The bankruptcy models in this investigation employed 19 financial variables (see Table 2), selected in previous research on financial distress, as input variables. These variables were organized into four groups, according to whether they related to liquidity, profitability, asset management or financial structure. The input variables of all the models are the same. The hit ratio of classification is the indicator used to evaluate the predictive accuracy of model. The bootstrap technique has been widely used in financial research to evaluate the external validity of model in prediction. Additional evidence of the stationary of the models was obtained by other samples. Thus, this study not only evaluates the within-sample predictive capacity (internal validity) but also employs the bootstrap technique to evaluate the predictive ability in simulated sample (external validity). Table 4 describes the matched sample in this study. A total of listed 88 firms were obtained from the literature on financial distress in Taiwan.

### 4.2. Normality test

Most multivariate models assume that the data are normally distributed. Thus, the normality of the input variables must be tested before these models can be applied. This study employed the Kolmogorov–Smirnov $Z$ test (KS $Z$) to test the data for 1 year prior to failure, 2 years prior to failure, and 3 years prior to fail, to determine the distribution of the experimental data. The result of normality test was illustrated in Table 5. As Table 5 shows, most of financial ratios were not normally distributed as has been stated in previous research.

### 4.3. Predictive accuracy of matched samples

The average predictive accuracy of the failing company model is 92.61% in the first year before failure, 91.19% in the second year, and 83.81% in the third. With reference

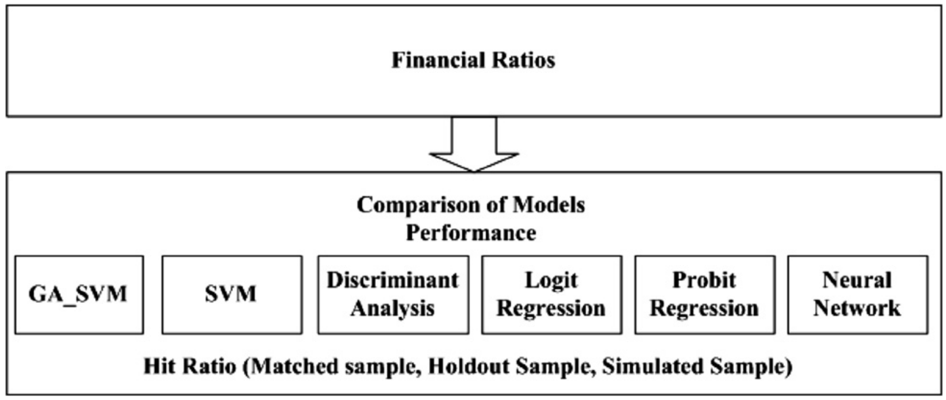


Fig. 2. Analysis steps.

Table 4
Description of the matched sample

| Industry | 1998 Number of firms | 1999 Number of firms | 2000 Number of firms | Total |
|---|---|---|---|---|
| Foods | 8 | 4 | 4 | 16 |
| Plastics | 0 | 12 | 0 | 12 |
| Fiber and textile | 0 | 8 | 0 | 8 |
| Electric, engineering, and machine | 4 | 0 | 0 | 4 |
| Electric wire and cable | 0 | 4 | 0 | 4 |
| Glass and ceramic | 0 | 4 | 0 | 4 |
| Steel | 4 | 8 | 4 | 16 |
| Motor | 4 | 0 | 0 | 4 |
| Electronics | 0 | 8 | 0 | 8 |
| Construction | 4 | 4 | 0 | 8 |
| Metal | 4 | 0 | 0 | 4 |
| Total | 28 | 52 | 8 | 88 |

Table 5
Normality test

| Ratio | Year – 1 KS $Z$ | Year – 2 KS $Z$ | Year – 3 KS $Z$ |
|---|---|---|---|
| *Liquidity* | | | |
| Current ratio | $2.120^{***}$ | $1.816^{***}$ | $1.546^{*}$ |
| Quick ratio | $2.326^{***}$ | $1.941^{***}$ | $1.803^{**}$ |
| Cash flow ratio | 1.155 | 1.003 | $1.899^{**}$ |
| *Profitability* | | | |
| Net income to sales | $2.735^{***}$ | $2.689^{***}$ | $2.935^{***}$ |
| Gross profit to sales | 0.579 | 1.046 | 1.046 |
| Net income to total assets | $1.976^{***}$ | $2.494^{***}$ | $2.747^{***}$ |
| Net income to stockholder's equity | $3.522^{***}$ | 1.074 | $2.466^{***}$ |
| Operating income to sales | 1.260 | 1.339 | 1.080 |
| Earning per share (EPS) | $2.477^{***}$ | 1.344 | $3.822^{***}$ |
| Growth ratio of sales | $1.824^{***}$ | $1.835^{***}$ | 1.147 |
| *Asset management* | | | |
| Total asset turnover | 1.190 | 1.387 | $1.457^{*}$ |
| Fixed assets turnover | $2.157^{***}$ | $2.212^{***}$ | $1.781^{**}$ |
| Inventory turnover | 1.508 | $1.783^{**}$ | $1.370^{*}$ |
| Receivables turnover | $2.252^{***}$ | $2.096^{***}$ | $2.190^{***}$ |
| *Financial structure* | | | |
| Debt ratio | 0.816 | 0.500 | 0.606 |
| Long-term liabilities to fixed assets | $2.285^{***}$ | $2.840^{***}$ | $2.353^{***}$ |
| Degree of financial leverage (DFL) | $4.142^{***}$ | $4.273^{***}$ | $4.058^{**}$ |
| Liabilities to stockholder's equity | $2.831^{***}$ | $1.476^{*}$ | $4.466^{***}$ |
| Interest coverage ratio | $4.211^{***}$ | $4.187^{***}$ | $4.504^{***}$ |

KS $Z$ denotes Kolmogorov–Smirnov $Z$ test.
$^{*}$ Denotes $\alpha < 0.1$.
$^{**}$ Denotes $\alpha < 0.05$.
$^{***}$ Denotes $\alpha < 0.01$.

to the first year before failure, the failed of firms predicted to non-failure (Type I error) is rarer than the non-failed of firms predicted to failure (Type II error). Artificial intelligence models (NN, SVM, and GA-SVM) are able to perfectly predict bankruptcy (100% accuracy), in the first, second and third years before failure. DA exhibited the lowest predictive accuracy of all the models. DA and probit model yielded the highest Type II and Type I errors, respectively, in all years before failure (Table 6).

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 the first year before failure, predictions of failed firms not to fail (Type I error) were greater than predictions of non-failed companies to fail (Type II error). In the first year before failure, the average Type I and II errors are 4.9% and 2.6%, respectively. In the second year before failure, the average Type I and II errors are 6.5% and 2.3%, respectively. In the third year before failure, the average Type I and II errors are 11.4% and 4.8%, respectively. Probit model and logit model yield the highest Type I error, while DA has the highest Type II error.

### 4.4. Predictive accuracy of holdout sample

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. Two thirds of the data are commonly used as the

Table 6
Accuracies of models on financial distress prediction

| Models | Year − 1 | | | Year − 2 | | | Year − 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | Accuracy | Type I error | Type II error | Accuracy | Type I error | Type II error | Accuracy | Type I error | Type II error |
| DA | 0.8750 | 0.023 | 0.102 | 0.8522 | 0.057 | 0.091 | 0.7500 | 0.080 | 0.170 |
| Logit | 0.9205 | 0.080 | 0.000 | 0.8978 | 0.102 | 0.000 | 0.8068 | 0.182 | 0.011 |
| Probit | 0.9090 | 0.091 | 0.000 | 0.8977 | 0.102 | 0.000 | 0.7955 | 0.193 | 0.011 |
| NN | 1.0000 | 0.000 | 0.000 | 1.0000 | 0.000 | 0.000 | 1.0000 | 0.000 | 0.000 |
| SVM | 1.0000 | 0.000 | 0.000 | 1.0000 | 0.000 | 0.000 | 1.0000 | 0.000 | 0.000 |
| GA-SVM | 1.0000 | 0.000 | 0.000 | 1.0000 | 0.000 | 0.000 | 1.0000 | 0.000 | 0.000 |
| Average | 0.9261 | 0.049 | 0.026 | 0.9119 | 0.065 | 0.023 | 0.8381 | 0.114 | 0.048 |

*Note:* DA denotes discriminant analysis; NN denotes neural network.
SVM refers to a support vector machine with fixed values of parameters.
GA-SVM denotes a support vector machine with values of parameters optimized by RGA.
Type I error represents the probability of misclassifying a failed firm.
Type II error represents the probability of misclassifying a non-failed firm.

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. In this study, the holdout sample is the group of data unused when the financial distress models are computed. This sample is used to validate applicability of the financial bankruptcy models to a separate sample of respondents, also called the validation sample (Hair, Anderson, Tatham, & Black, 1998). The holdout sample herein consists of all firms listed on the TSE and OTC in 2001 and 2002. The holdout sample therefore includes four data files. The file (2001TSE.txt) contained 605 records but 67 records were deleted because data were missing. For the same reason, 69 records were removed from the file (2002TSE.txt). The file (2001OTC) included 373 records, after 33 of the original 406 records were removed and file (2002OTC) included 356 records, after 50 of the original 356 records were removed. The models used to predict financial distress were trained using the preceding year's data. Results revealed that the proposed model, GA-SVM, outperformed the other bankruptcy models.

As Table 7 shows, the average predictive accuracy was 87.33% in predicting companies in the TSE market that failed in 2001, and 74.9% in 2002. The proposed model, GA-SVM, outperformed other bankruptcy models in 2001 and 2002 years. The GA-SVM had the highest predictive accuracy, the highest Type I error and the lowest Type II error. DA had the worst predictive accuracy of all models but it had the lowest Type I error. Table 8 presents results concerning the prediction of bankruptcy in the OTC market in 2001–2002. The average predictive accuracy is 79.50% in predicting companies in the OTC market that failed in 2001, and 78.47% in 2002. The financial distress model performed better for the OTC market than the TSE market in 2001–2002. The overall predictive accuracies exceeded that for the TSE market. The GA-SVM still outperformed among the other financial distress models in 2001 and 2002 years. The GA-SVM exhibited the highest predictive accuracy, the lowest Type I and Type II errors for the 2001 OTC market. DA and probit model had the lowest predictive accuracy for the OTC market in both 2001 and 2002.

### 4.5. Predictive accuracy of simulated sample

The bootstrap technique was introduced by Efron and is fully elucidated in Efron and Tibshirani (1993). Given a dataset of size $n$, a bootstrap sample is constructed by

Table 7
Predictive accuracies of models in holdout sample (the TSE market)

| Models | 2001 TSE (538 firms)<br>Year + 1 | | | 2002 TSE (534 firms)<br>Year + 2 | | |
|---|---|---|---|---|---|---|
| | Accuracy | Type I error | Type II error | Accuracy | Type I error | Type II error |
| DA | 0.7937 | 0.000 | 0.206 | 0.7846 | 0.009 | 0.206 |
| Logit | 0.8773 | 0.007 | 0.115 | 0.6985 | 0.015 | 0.287 |
| Probit | 0.8922 | 0.004 | 0.104 | 0.5524 | 0.019 | 0.429 |
| NN | 0.8271 | 0.006 | 0.167 | 0.5131 | 0.017 | 0.470 |
| SVM | 0.8662 | 0.006 | 0.128 | 0.9738 | 0.026 | 0.000 |
| GA-SVM | 0.9833 | 0.013 | 0.004 | 0.9738 | 0.026 | 0.000 |
| Average | 0.8733 | 0.006 | 0.121 | 0.7494 | 0.019 | 0.232 |

*Note:* DA denotes discriminant analysis; NN denotes neural network.
SVM refers to a support vector machine with fixed values of parameters.
GA-SVM denotes a support vector machine with values of parameters optimized by RGA.
Type I error represents the probability of misclassifying a failed firm.
Type II error represents the probability of misclassifying a non-failed firm.

Table 8
Predictive accuracies of models in holdout sample (the OTC market)

| Models | 2001 OTC (373 firms) Year + 1 | | | 2002 OTC (356 firms) Year + 2 | | |
|---|---|---|---|---|---|---|
| | Accuracy | Type I error | Type II error | Accuracy | Type I error | Type II error |
| DA | 0.5899 | 0.000 | 0.410 | 0.7416 | 0.006 | 0.253 |
| Logit | 0.7837 | 0.000 | 0.216 | 0.6067 | 0.011 | 0.382 |
| Probit | 0.6938 | 0.000 | 0.306 | 0.5056 | 0.011 | 0.483 |
| NN | 0.7135 | 0.000 | 0.287 | 0.8820 | 0.014 | 0.104 |
| SVM | 0.9944 | 0.000 | 0.006 | 0.9860 | 0.014 | 0.000 |
| GA-SVM | 0.9944 | 0.000 | 0.006 | 0.9860 | 0.014 | 0.000 |
| Average | 0.7950 | 0.000 | 0.2051 | 0.7847 | 0.012 | 0.204 |

randomly sampling *n* instances uniformly from the original data. In this section, the simulated sample is constructed by bootstrap techniques. The original sample of 88 enterprises is divided into a training sample of 60 enterprises, and the remaining 28 enterprises are used as the validation sample. The ratio of the size of the training sample to that of the validation sample is designed to be 2:1. The financial distress models are predicting for varying samples, by bootstrapping from 50 to 1000 times, to evaluate the reliability of validation.

Table 9
Predictive accuracies of bankruptcy models for simulated samples

| Bootstrap times | Predictive accuracy of bankruptcy models | | | | | |
|---|---|---|---|---|---|---|
| | DA | Logit | Probit | NN | Fix-SVM* | GA-SVM |
| 50 | 0.68 | 0.69 | 0.69 | 0.69 | 0.64 | 0.72 |
| 100 | 0.71 | 0.69 | 0.70 | 0.68 | 0.66 | 0.75 |
| 200 | 0.71 | 0.69 | 0.69 | 0.70 | 0.65 | 0.75 |
| 300 | 0.72 | 0.69 | 0.69 | 0.70 | 0.64 | 0.75 |
| 400 | 0.72 | 0.70 | 0.71 | 0.70 | 0.65 | 0.76 |
| 500 | 0.71 | 0.69 | 0.69 | 0.70 | 0.65 | 0.76 |
| 1000 | 0.71 | 0.69 | 0.69 | 0.70 | 0.65 | 0.75 |

*Note:* Predictive accuracies of bankruptcy models are represented as percentages.
* Fix-SVM denotes that SVM is run with fix values of SVM parameters.

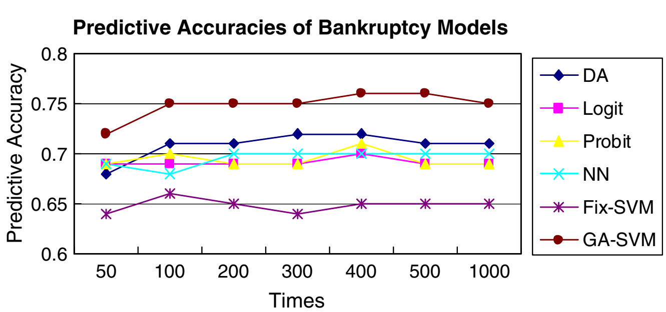


Fig. 3. Predictive accuracies of bankruptcy models for simulated sample.

Table 9 and Fig. 3 present and plot, respectively, the predictive accuracies of the bankruptcy models obtained after bootstrapping various numbers of times. As results shown, GA-SVM performed well when applied to the simulated sample when bootstrapping was performed various numbers of times. Traditional SVM with fixed values of parameters had the lowest predictive accuracy, regardless of the number of bootstrapping times. The results implied that the predictive accuracies of the SVM were dramatically increased by using the optimal parameters. Additionally, GA-SVM and SVM run with fixed values (Fix-SVM) have the lowest Type I error, but that the Fix-SVM exhibited the highest Type II error for simulated sample (Table 10).

## 5. Discussions

This study develops a novel model to search the optimal values of SVM parameters, to increase the accuracy of prediction and ability of generalization. DA, logit, probit, NN, SVM, and the proposed model (GA-SVM) were applied to a dataset on bankruptcies in Taiwan. First, this study found that the RGA yields different optimal values of the parameters of SVM given various datasets (paired sam-

Table 10
Type I and II errors of bankruptcy models for simulated samples

| Bootstrap times | Type I error % | | | | | | Type II error % | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | DA | Logit | Probit | NN | Fix SVM | GA SVM | DA | Logit | Probit | NN | Fix SVM | GA SVM |
| 50 | 0.19 | 0.18 | 0.19 | 0.17 | 0.08 | 0.13 | 0.12 | 0.13 | 0.12 | 0.14 | 0.27 | 0.15 |
| 100 | 0.16 | 0.18 | 0.19 | 0.17 | 0.11 | 0.13 | 0.13 | 0.13 | 0.12 | 0.15 | 0.23 | 0.12 |
| 200 | 0.16 | 0.18 | 0.19 | 0.17 | 0.11 | 0.12 | 0.12 | 0.13 | 0.12 | 0.12 | 0.23 | 0.12 |
| 300 | 0.16 | 0.18 | 0.19 | 0.16 | 0.10 | 0.12 | 0.12 | 0.14 | 0.13 | 0.14 | 0.26 | 0.13 |
| 400 | 0.16 | 0.17 | 0.18 | 0.16 | 0.11 | 0.12 | 0.12 | 0.12 | 0.11 | 0.14 | 0.24 | 0.13 |
| 500 | 0.16 | 0.17 | 0.19 | 0.17 | 0.11 | 0.12 | 0.13 | 0.14 | 0.13 | 0.14 | 0.24 | 0.12 |
| 1000 | 0.17 | 0.18 | 0.19 | 0.17 | 0.11 | 0.12 | 0.13 | 0.14 | 0.13 | 0.14 | 0.24 | 0.13 |

*Note:* Predictive accuracies of bankruptcy models are represented percentages.

ple, holdout sample, and simulated sample). The optimal values of SVM parameters are not constant but fall in the range and the range of each parameter differs with the dataset. This work found that the optimal values of the two parameters of SVM differed with the dataset, and changed at each time. The results reveal that the optimal values have a range, and are not constants which provide a direction for future investigation. Secondly, properly determining the values of SVM parameters were drastically increased the accuracy of the prediction of bankruptcies. The optimal kernel parameters of SVM can be automatically determined via the proposed approach.

Thirdly, artificially intelligent models (GA-SVM, SVM, and NN) are more accurate in predicting financial distress than other multivariate statistical models. Experimental results show that the GA-SVM model performs the best, implying that the hybrid system has a high potential to dramatically increase the predictive accuracy when integrating GA with traditional SVM model. Fourthly, most financial ratios did not satisfy the normality assumption for multivariate statistical models such as the MDA and the probit model. Thus, MDA exhibited the worst predictive accuracy and the largest errors of all models tested herein. Finally, the structural risk minimization principle (SRM) has been shown to be superior to traditional empirical risk minimization principle (ERM) which employed by conventional neural networks, was embodied in SVM. SRM is able to minimize an upper bound of generalization error as opposed to ERM that minimizes the error on training data. Thus, the solution of SVM may be global optimum while other neural network models tend to fall into a local optimal solution, and overfitting is unlikely to occur with SVM (Cristianini et al., 1998; Hearst, Dumais, Osuna, Platt, & Scholkopf, 1998; Kim, 2003). Based on these reasons, we can conclude that the proposed model more accurately predicts bankruptcy than the other tested models of bankruptcy. Additionally, the results of this work demonstrate that the predictive accuracy of the SVM in forecasting the financial distress of corporations is significantly increased by optimizing its parameters.

## 6. Conclusions

This study pioneered on applying the GA-SVM to financial distress prediction. Therefore, the primary goal of this study is to apply this new model to increase the predictive accuracy of financial failure. Empirical results reveal that the proposed GA-SVM model is a very promising hybrid SVM model for predicting bankruptcy in terms of both predictive accuracy and generalization ability. The proposed GA-SVM model can be automated to determine the optimal values of SVM parameters and exhibits increased predictive accuracy in given various datasets. Grice and Ingram (2001) reported that Altman's $Z$-score model declined when applied to various industries. In other words, the $Z$-score model was sensitive to industry classifications. In addition, both theories and experiments have shown that benefits do not always accrue as the computational cost is increased, especially if the relative accuracies are more important that the exact values. The results imply that the models with high predictive accuracies in sample do not guarantee to the same high accuracies in holdout sample prediction. The contribution of this study demonstrated that the proposed model performed well when applied in the holdout sample, revealing the generalizability of this model to forecast financial distressed firms in various industries. Thus, the forecasting technique (GA-SVM) can be developed and coded as a commercial package for investors. The result of the GA-SVM (financial distress pre-warning system) can provide a guide of investment for investors and government.

## References


Adewuya, A. A. (1996). *New methods in genetic search with real-valued chromosomes*. Master's thesis, Cambridge: Massachusetts Institute of Technology.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*(4), 589–609.

Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). ZETA™ analysis—A new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance, 1*(1), 29–54.

Beaver, W. H. (1966). Financial ratios as predictors of failures. *Journal of Accounting Research, 4*(3), 71–102 (Suppl.).

Blum, M. (1974). Failing company discriminant analysis. *Journal of Accounting Research, 12*(1), 1–25 (Spring).

Boster, B., Guyon, I., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on computational learning theory* (pp. 144–152).

Campbell, C. (2002). Kernel methods: a survey of current techniques. *Neurocomputing, 48*(1–4), 63–84.

Cao, L. J. (2003). Support vector machines experts for time series forecasting. *Neurocomputing, 51*(1–4), 321–339.

Claessens, S., Djankov, S., & Klapper, L. (2003). Resolution of corporate distress in East Asia. *Journal of Empirical Finance, 10*(1/2), 199–216.

Cortes, C., & Vapnik, V. N. (1995). Support vector networks. *Machine Learning, 20*(3), 273–297.

Cristianini, N., Shawe-Taylor, J., & Campell, C. (1998). Dynamically adapting kernels in support vector machines. *Advances in Neural Information Processing Systems, 11*, 204–210.

Deakin, E. B. (1972). A discriminant analysis of predictors of business failure. *Journal of Accounting Research, 10*(1), 167–179 (Spring).

Duan, K., Keerthi, S. S., & Poo, A. N. (2003). Evaluation of simple performance measures for tuning SVM hyperparameters. *Neurocomputing, 51*(1–4), 41–59.

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. Chapman and Hall.

Fogel, D. B. (1994). An introduction to simulated evolutionary optimization. *IEEE Transactions on Neural Networks, 5*(1), 3–14.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization & machine learning*. Reading, MA: Addision-Wesley.

Grefenstette, J. J. (1986). Optimization of control parameters for genetic algorithms. *IEEE Transactions on System, Man and Cybernetics, 16*(1), 122–128.

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

Hair, J. F., Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Prentice-Hall International Inc.

Haupt, R. L., & Haupt, S. E. (1998). *Practical genetic algorithms*. Wiley Interscience Publication.

Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. *IEEE Expert Intelligent Systems and their Applications, 13*(4), 18–28.

Huang, Y. P., & Huang, C. H. (1997). Real-valued genetic algorithms for fuzzy grey prediction system. *Fuzzy Sets and Systems, 87*(3), 265–276.

Kecman, V. (2001). *Learning and soft computing: Support vector machines, neural networks, and fuzzy logic models*. Cambridge, MA, London, England: MIT Press.

Kim, K. J. (2003). Financial time series forecasting using support vector machines. *Neurocomputing, 55*(1/2), 307–320.

Lee, K. C., Han, I., & Kwon, Y. (1996). Hybrid neural network models for bankruptcy predictions. *Decision Support Systems, 18*(1), 63–72.

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

Lin, P. T. (2001). Support vector regression: Systematic design and performance analysis, Unpublished Doctoral Dissertation, Department of Electronic Engineering, National Taiwan University.

Mattera, D., & Haykin, S. (1999). Support vector machines for dynamic reconstruction of a chaotic system. In B. Schölkopf, C. J. C. Burges, & A. J. Smola (Eds.), *Advances in kernel methods – Support vector learning* (pp. 211–242). Cambridge, MA: MIT Press.

McCall, J., & Petrovski, A. (1999). A decision support system for cancer chemotherapy using genetic algorithms. In *Proceedings of the international conference on computational intelligence for modeling, control and automation* (pp. 65–70).

Min, J. H., & Lee, Y. C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications, 28*(4), 603–614.

Mukherjee, S., Osuna, E., & Girosi, F. (1997). Nonlinear prediction of chaotic time series using a support vector machine. In *NNSP'97: Neural networks for signal processing VII: Proceedings of the IEEE signal processing society workshop*. Amelia Island, FL, USA (pp. 511–520).

Müller, K. R., Smola, A. J., Rätsch, G., & Schölkopf, K. J. (1999). Using support vector machines for time series prediction. In B. Scholkopf, C. J. C. Burges, & A. J. Smola (Eds.), *Advances in kernel methods—support vector learning* (pp. 243–254). Cambridge, MA: MIT Press.

Odom, M., & Sharda, R. (1990). A neural network for bankruptcy prediction. In *Proceedings of the IEEE international conference on neural network*, vol. 2 (pp. 163–168).

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*(1), 109–131.

Pelckmans, K., Suykens, J. A. K., Van Gestel, T., Brabanter, J. De, Lukas, L., Hamers, B., Moor, B. De, & Vandewalle, J. (2002). LS-SVM Toolbox User's Guide Version 1.4, November.

Platt, H. D., & Platt, M. B. (1990). Development of a class of stable predictive variables: the case of bankruptcy prediction. *Journal of Business Finance and Accounting, 17*(1), 31–51 (Spring).

Platt, H. D., & Platt, M. B. (2002). Predicting corporate financial distress: reflections on choice-based sample bias. *Journal of Economics and Finance, 26*(2), 184–199.

Salcedo-Sanz, S., Fernandez-Villacanas, J. L., Segovia-Vargas, M. J., & Bousono-Calzon, C. (2005). Genetic programming for the prediction of insolvency in non-life insurance companies. *Computers and Operations Research, 32*(4), 749–765.

Su, W. C. (2000). An empirical study of bankruptcy prediction on listed companies in Taiwan. Unpublished Master thesis, Department of International Business, National Dong Hwa University (in Chinese).

Tam, K. Y. (1991). Neural network models and the prediction of bankruptcy. *OMEGA: The International Journal of Management Science, 19*(5), 429–445.

Tam, K. Y., & Kiang, M. Y. (1992). Managerial applications of neural networks: the case of bank failure predictions. *Management Science, 38*(7), 926–947.

Tay, F. E. H., & Cao, L. J. (2001). Improved financial time series forecasting by combining support vector machines with self-organizing feature map. *Intelligent Data Analysis, 5*(4), 339–355.

Tay, F. E. H., & Cao, L. J. (2002). Modified support vector machines in financial time series forecasting. *Neurocomputing, 48*(1–4), 847–861.

Tseng, F. M., & Lin, L. (2005). A quadratic interval logit model for forecasting bankruptcy. *OMEGA: The International Journal of Management Science, 33*(1), 85–91.

Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York: Springer.

Wu, C. H. (2004). A new application of support vector machine to the prediction of corporate failure. In *Proceeding of international conference of Pacific RIM management*.

Yang, Z. R., Platt, M. B., & Platt, H. D. (1999). Probabilistic neural networks in bankruptcy prediction. *Journal of Business Research, 44*(2), 67–74.

Zavgren, C. (1985). Assessing the vulnerability to failure of American industrial firms: a logistic analysis. *Journal of Business, Finance and Accounting, 12*(1), 19–45.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research, 22*, 59–82 (Suppl.).