A HYBRID MODEL FOR BUSINESS FAILURE PREDICTION – UTILIZATION OF PARTICLE SWARM OPTIMIZATION AND SUPPORT VECTOR MACHINES

*Mu-Yen Chen**

Abstract: Bankruptcy has long been an important topic in finance and accounting research. Recent headline bankruptcies have included Enron, Fannie Mae, Freddie Mac, Washington Mutual, Merrill Lynch, and Lehman Brothers. These bankruptcies and their financial fallout have become a serious public concern due to huge influence these companies play in the real economy. Many researchers began investigating bankruptcy predictions back in the early 1970s. However, until recently, most research used prediction models based on traditional statistics. In recent years, however, newly-developed data mining techniques have been applied to various fields, including performance prediction systems. This research applies particle swarm optimization (PSO) to obtain suitable parameter settings for a support vector machine (SVM) model and to select a subset of beneficial features without reducing the classification accuracy rate. Experiments were conducted on an initial sample of 80 electronic companies listed on the Taiwan Stock Exchange Corporation (TSEC).

This paper makes four critical contributions: (1) The results indicate the business cycle factor mainly affects financial prediction performance and has a greater influence than financial ratios. (2) The closer we get to the actual occurrence of financial distress, the higher the accuracy obtained both with and without feature selection under the business cycle approach. For example, PSO-SVM without feature selection provides 89.37% average correct cross-validation for two quarters prior to the occurrence of financial distress. (3) Our empirical results show that PSO integrated with SVM provides better classification accuracy than the Grid search, and genetic algorithm (GA) with SVM approaches for companies as normal or under threat. (4) The PSO-SVM model also provides better prediction accuracy than do the Grid-SVM, GA-SVM, SVM, SOM, and SVR-SOM approaches for seven well-known UCI datasets. Therefore, this paper proposes that the PSO-SVM approach could be a more suitable method for predicting potential financial distress.

Key words: *Particle swarm optimization, support vector machine, business failure prediction*

*Mu-Yen Chen
Department of Information Management, National Taichung Institute of Technology, Taichung 404, Taiwan, R.O.C, E-mail: `mychen@ntit.edu.tw`

Received: October 1, 2010
Revised and accepted: February 15, 2011

1. Introduction

The Asian financial crisis started on July 2, 1997 with a 15-20% devaluation of the Thai Baht, after two months of massive speculative attacks and a little more than a month after the bankruptcy of Thailand's largest finance company, *Finance One* [1]. The crisis resulted in unprecedented economic and financial hardship, which spread to Taiwan as well. Taiwan's export growth rate dropped from 5.3% in 1997 to −9.4% the following year, coinciding with a drop in industrial production from 7.4% to 2.6% [2]. The TAIEX (Taiwan Stock Exchange Capitalization Weighted Stock Index) peaked at 10,116 on August 26, 1997, but the ensuing flight of foreign capital and crash in investor confidence caused the TAIEX to drop to a low of 5,474 by February 1999. This caused a major upheaval in Taiwan's financial markets, with many investors incurring heavy losses. More recently, the crisis of 2007–2010 has proven to be the worst financial calamity since the Great Depression of the 1930s [3]. Housing bubbles, credit booms, sub-prime and predatory lending, incorrect pricing of risk and the collapse of the shadow banking system dealt a heavy economic shock to the U.S. economy. Several major institutions, including Lehman Brothers, Merrill Lynch, Fannie Mae, Freddie Mac, Washington Mutual, Wachovia, and AIG, either failed, were acquired under duress, or were taken over by the government [4]. The FDIC (Federal Deposit Insurance Corporation) announced 111 U.S. bank failures from January to August 2010, and predicted another 140 failures in the remainder of 2010 [5]. This crisis rapidly spread from the U.S. to the global economy, putting pressure on all major sources of external revenue for developing countries. Moreover, Taiwan's electronics industry, boasting 670 companies listed on the Taiwan Stock Exchange (TSE) and the OTC (Over-the-Counter) Securities Market and having an annual production value of US$300 billion, will play the role of locomotive in Taiwan's long-term and world economic development. Thus, despite the existence of several methods for predicting corporate failure, it is worthwhile for researchers and industry operators to continuously develop state-of-the-art methods reflecting various symptoms of corporate failure that may not be explained by the existing methods [6]. Therefore, we will apply evolutionary techniques into financial bankruptcy prediction system construction with Taiwan's electronic industry.

Since the late 1960s considerable research has been conducted with the goal of accurately predicting the failure of financial firms. Initial approaches were based on statistical methodologies pioneered by Beaver (1966) who used univariate analysis to build a financial prediction model for banks [7]. Altman (1968) later pointed out drawbacks in Beaver's model, and used discriminant analysis (DA) to rebuild the model [8]. Several years later, Martin (1977) developed the stochastic model with logistic regression to measure the probability of bank failure based on data obtained from the Federal Reserve System [9]. In 1984, Zmijewski (1984) proposed another stochastic model with probit analysis to weight the log-likelihood function by the ratio of the population frequency rate to the sample frequency rate of individual groups, bankrupt and non-bankrupt [10]. Cielen, Peeters, and Vanhoof (2004) [11] and Kao and Liu (2004) [12] used data envelopment analysis (DEA) to predict the bank failures but strict assumptions, such as linearity and normality, and independence among predictable variables limits the applicability of their models in real world. Yet another approach employed artificial intelligence (AI) methods and, beginning in the 2000s, a number of studies have applied these methods to bankruptcy prediction problems.

Hu and Tseng (2007) integrated fuzzy integral theory with genetic algorithms (GA) to predict bankruptcy [13], and their proposed method performed well in comparison with traditional functional-link net and multivariate techniques. Sanchis et al. (2007) presented a new approach to predicting financial stability and insurance insolvency using rough sets [14], with quite satisfactory results. Kirkos et al. (2007) investigated the prediction accuracy of Decision Trees (DT), Artificial Neural Networks (ANN) and Bayesian Belief Networks (BBN) in the identification of fraudulent financial statements [15]. The results showed that BBN provided more accurate predictions than ANN and DT. Sun and Li (2008) designed an entropy-based discretization method for predicting financial distress in listed companies [16]. An empirical experiment with 35 financial ratios and 135 pairs of listed companies indicated that the DT model constructed by the data mining method provided satisfactory prediction accuracy in both the training and validation samples. Ahn and Kim (2009) integrated GA and case-based reasoning (CBR) to predict bankruptcy crises [17], with experimental results indicating their proposed model can significantly improve the predictive accuracy of conventional CBR. Chen and Du (2009) compared the predictive performance of ANN with traditional statistical and data mining techniques for listed companies in Taiwan [18], finding that the ANN approach provided greater accuracy than the traditional statistical and DM clustering approaches. Boyacioglu et al. (2009) applied various neural network techniques, support vector machines and multivariate statistical methods to predict the failure of Turkish banks [19]. The results showed that multi-layer perceptron (MLP) and learning vector quantization (LVQ) are potentially the most successful models in predicting bank failures.

However, there have been little researches into swarm-inspired optimization techniques for bankruptcy prediction. Since most real world problems are multi-criteria problems, it would seem appropriate to use multi-objective algorithms in seeking solutions. Therefore, this paper aims to effectively solve continuous financial datasets. We integrate a novel particle swarm optimization (PSO) algorithm with a support vector machines (SVM) classification model. The proposed algorithm can reduce the probability of being trapped in local optima and enhance accuracy and global search capabilities. The proposed PSO-SVM model will also be compared with grid search (Grid-SVM), GA-SVM, SVM, Self-Organizing Map (SOM), and Support Vector Regression (SVR) - SOM models. The main objectives of this paper are to (1) adopt swarm-inspired optimization techniques to construct a financial distress prediction model, (2) use financial ratios and business cycle index to improve the accuracy of the financial distress prediction model, (3) compare the accuracy of PSO-SVM and other neural networks approaches, and (4) to expand this model so that it will work within a financial distress prediction system as a type of early warning system.

The remainder of this paper is organized as follows: Section 2 provides an overview of SVM and PSO. Section 3 describes the PSO-SVM hybrid model. Section 4 presents the experimental results from a simulation dataset. Conclusions are presented in Section 5, along with recommendations for future research.

2. Literature Review

2.1 MLP and SVM neural networks

An MLP is a feed-forward artificial neural network model that maps sets of input data onto a set of appropriate outputs. For neural networks, MLP uses a back-propagation learning algorithm [20] as the standard algorithm for the subject of ongoing research in computational neuroscience and soft computing. Fig. 1 shows the l–m–n architecture of a MLP model (l denotes input neurons, m denotes hidden neurons, and n denotes output neurons). Neurons in the input layer have a pure linear activation function, but some nonlinear activation functions, such as logarithmic and tangent sigmoid functions, are used in the neurons in hidden and output layers. The input signals are modified by the interconnection weight, known as weight factor w_{ji}, which represents the interconnection of the ith node of the first layer to the jth node of the second layer. The sum of the modified signals (total activation) is then modified by a sigmoid transfer function(f). Similarly, the output signals of the hidden layer are modified by interconnection weight w_{kj} of the kth node of the output layer to the jth node of the hidden layer. The sum of the modified signals is then modified using the sigmoid transfer (f) function and the output is collected at the output layer.

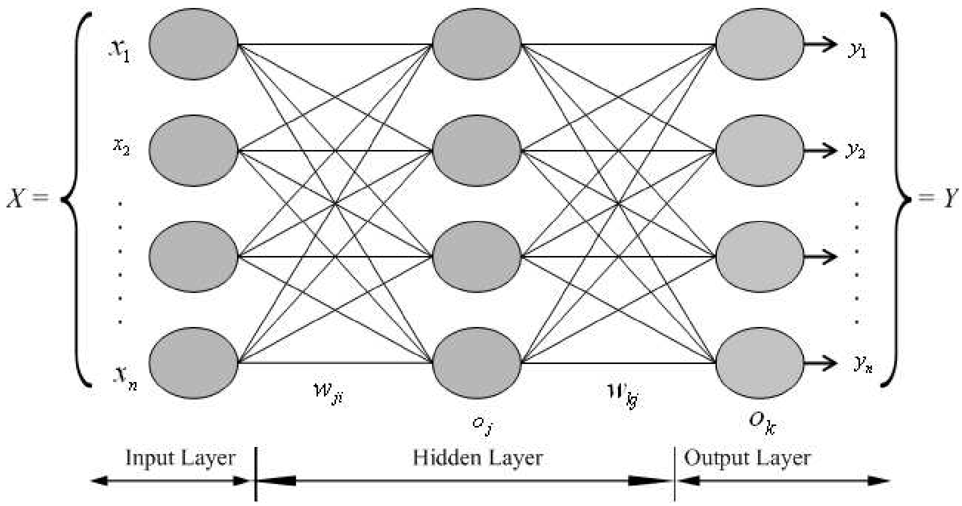

Fig. 1 *MLP architecture.*

Let $I_p = (I_{p1}, I_{p2}, \ldots I_{pl}), p = 1, 2 \ldots N$ be the pth pattern among N input patterns, where w_{ji} and w_{kj} are respectively the connection weights between the ith input neuron to the jth hidden neuron, and the jth hidden neuron to the kth output neuron [21].

Output from a neuron in the input layer is

$$O_{pi} = I_{pi}, \quad i = 1, 2, \ldots, l. \tag{1}$$

Output from a neuron in the hidden layer is,

$$O_{pj} = f\left(NET_{pj}\right) \quad = f\left(\sum_{i=1}^{l} w_{ji} o_{pi}\right), \text{ for } 1 \le j \le m \tag{2}$$

Output from a neuron in the output layer is,

$$O_{pk} = f\left(NET_{pk}\right) \quad = f\left(\sum_{j=1}^{m} w_{kj} o_{pj}\right), \text{ for } 1 \le k \le n, \tag{3}$$

where $f()$ is the sigmoid transfer function given by $f(x) = 1/(1 + e^{-x})$.

Given a set of training examples, each marked as belonging to one of two categories, Vapnik's Support Vector Machine (SVM) training algorithm [22] builds a model that predicts whether a new example falls into one category or in the other. More formally, an SVM constructs a hyperplane or a set of hyperplanes in a high or infinite dimensional space, which can then be used for classification, regression or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data points of any class (the so-called functional margin), since, in general, the larger the margin the lower the generalization error of the classifier. The SVM architecture and hyperplane representation is shown in Fig. 2.

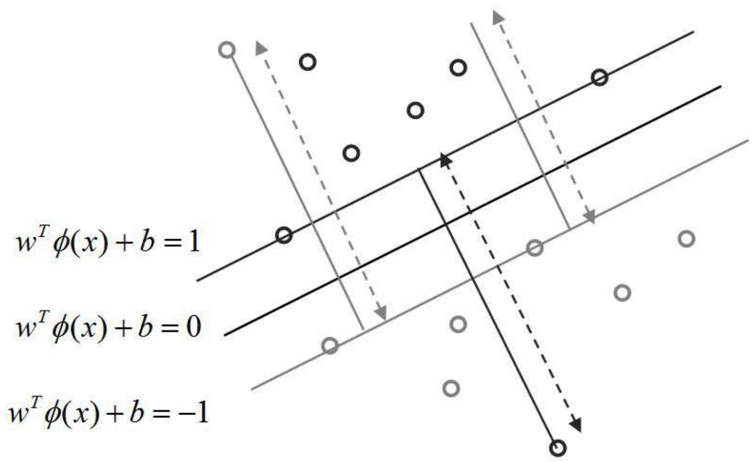

Fig. 2 *SVM architecture and hyperplane representation.*

In support vector machines, the input vector x is mapped to the high-dimensional feature space using the mapping function Φ(x) to enhance linear separability. Let the M training input–output pairs be $(x_i, y(x_i)), i = 1, \ldots, M$, where $y(x_i) = 1$ if x_i belongs to Class 1, and $y(x_i) = -1$ if x_i belongs to Class 2. If the training data are linearly separable in the feature space, we can obtain the decision function:

$$f(x) = w^T \phi(x) + b, \tag{4}$$

where w is a weight vector, b is a bias term, and $y(x_i) f(x_i) > 0$ for $i = 1, \ldots, M$. For unknown data x, if $f(x) \geq 0$, the data are classified into Class 1 and, if $f(x) < 0$, into Class 2. The distance between the separating hyperplane and the training datum nearest to the hyperplane is called the margin. The hyperplane with the maximum margin is called the optimal separating hyperplane that separates two classes. If the classification problem is not linearly separable in the feature space, the optimal separating hyperplane can be obtained by solving the following optimization problem:

$$\text{Minimize } Q(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{M} \zeta_i \tag{5}$$

$$\begin{aligned} &\text{Subject to } y(x_i)(w^T \phi(x) + b) \geq 1 - \zeta_i \\ &\qquad\qquad\qquad \text{For } i = \mathit{1},\ \mathit{2},\ \ldots,\ M \end{aligned} \tag{6}$$

where C is the regularization parameter that determines the trade off between the maximization of the margin and the minimization of the classification error, and ζ_i is the non-negative slack variable for x_i. Some of the kernels that are used in the support vector machines are as follows:

$K(x, x')$ is a kernel function that is given by

$$K(x, x') = \phi(x)^T \phi(x'). \tag{7}$$

The polynomial kernel with degree d is given by

$$K(x, x') = (x^T x' + 1)^d. \tag{8}$$

The radial basis function (RBF) kernel is given by

$$K(x, x') = \exp(-\gamma \|x - x'\|^2), \tag{9}$$

where γ is a positive parameter for slope control.

By introducing Lagrange multipliers α_i, the SVM training procedure amounts to solving a convex quadratic problem (QP). The solution is a unique globally-optimized result with the following properties

$$w = \sum_{i}^{N} \alpha_i y_i x_i. \tag{10}$$

Only if the corresponding $\alpha_i > 0$, are the x_i called support vectors. When the SVMs are trained, the decision function can be written as

$$f(x) = sign\left(\sum_{i=1}^{N} \alpha_i y_i (x \cdot x_i) + b\right) \tag{11}$$

In addition to applying SVMs to classification problems, Vapnik and his colleagues proposed a special version for regression, called SVR [23], entailing an ε − insensitive zone in the error loss function (see Fig. 2). This zone represents the degree of precision to which the bounds of generalization ability apply. Training vectors that lie within this zone are deemed correct, whereas those outside this zone are deemed incorrect, and contribute to the error loss function. These incorrect vectors become the support vectors. Such loss functions usually lead to a sparse representation of the decision rule, resulting in significant algorithmic and representational advantages. Besides, the SVR still contains all the main features that characterize the maximum margin algorithm: a non-linear function is learned by linear learning machine mapping onto a high-dimensional kernel-induced feature space. Moreover, one of the advantages of the SVR is that it can be used to avoid the difficulties inherent in linear functions in high-dimensional feature space, and the optimization problem is transformed into a dual convex QP [24].

2.2 Particle swarm optimization

The PSO algorithm was first introduced by Eberhart and Kennedy (1995) [25]. PSO is designed to simulate social behavior. Like evolutionary algorithms, PSO executes searches using a population (called a *swarm*) of individuals (called *particles*) that are renewed from iteration to iteration. All the particles have fitness values that are evaluated according to the fitness function to be optimized, and have velocities that direct the flight of the particles. It searches for the optimal value by sharing historical information and social information between the individual particles [26]. Moreover, through cooperation and competition among the population, population-based optimization approaches can often arrive at very good solutions efficiently and effectively. The advantages of PSO are that PSO is easy to implement in a few lines of computer code, it has few parameters requiring adjustment, and it is computationally inexpensive in terms of memory requirements and run time. Early testing has found the implementation to be effective with several kinds of problems, such as function optimization, artificial neural network training, fuzzy system control, and other areas [27].

A particle represents a potential problem solution move through a d-dimensional search space. Each particle i represents a candidate position, remembering the best value and the current position which had resulted in that value, called *pbest.* When a particle takes the entire population as its topological neighbors, the best value is a global best and is called *gbest.* All particles can share information about the search space. The d-dimensional position for particle i at iteration t can be represented as $x_i^t = \{x_{i1}^t, x_{i2}^t, \ldots, x_{id}^t\}$. Likewise, the velocity, which is also a d-dimensional vector, for particle i at iteration t can be described as $v_i^t = \{v_{i1}^t, v_{i2}^t, \ldots, v_{id}^t\}$. Fig. 3 illustrates the concept of modulating searching points. Let P_{id} denote the best previous position encountered by the ith particle. P_{gd} denotes the global best position thus far. The current velocity of the dth dimension of the ith particle at iteration t is as follows:

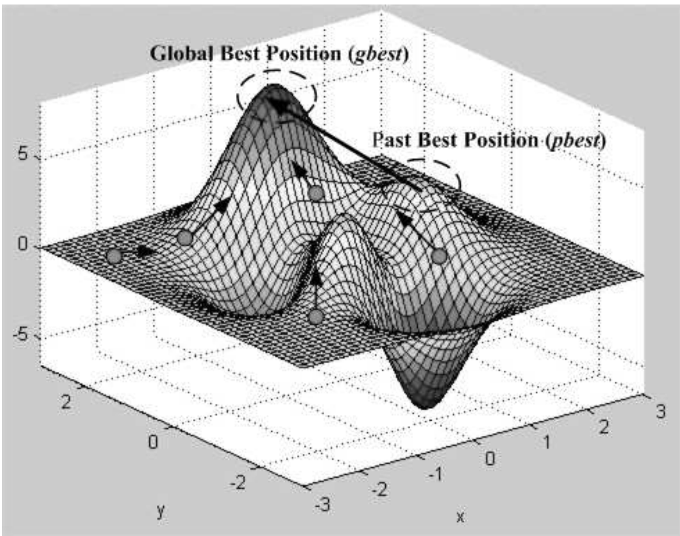

Fig. 3 *Search concept of particle swarm optimization.*

$$V_{id}^{t} = wV_{id}^{t-1} + c_1 r_1 (Pid^{t} - x_{id}^{t}) + c_2 r_2 (P_{gd}^{t} - x_{id}^{t}), \quad d = 1, 2, \ldots, D. \tag{12}$$

In the above formula, $r()$ is a random function in the range [0, 1], positive constants c_1 and c_2 are personal and social learning factors, and w is the inertia weight [28]. The velocity is restricted to the $[-V_{max}, V_{max}]$ range in which V_{max} is a predefined boundary value. The new position of a particle is calculated using the following formula:

$$X_{id}^{t} = X_{id}^{t} + V_{id}^{t}, d = 1, 2, \ldots, D \tag{13}$$

Unlike in GA, evolutionary programming and evolution strategies, in PSO the selection operation is not performed [29]. All particles in PSO are kept as members of the population through the course run. The velocity of the particle is updated according to its own previous best position and the previous best positions of its companions. The particles then fly with the updated velocities. PSO is the only evolutionary algorithm that does not implement survival of the fittest [30].

3. Research Methodology and Materials

3.1 Research methodology

In this research, we integrated the PSO and SVM techniques to create an early warning evaluation model of firms' financial structures. The research methodology is illustrated in Fig. 4. The three steps for building a financial bankruptcy crisis

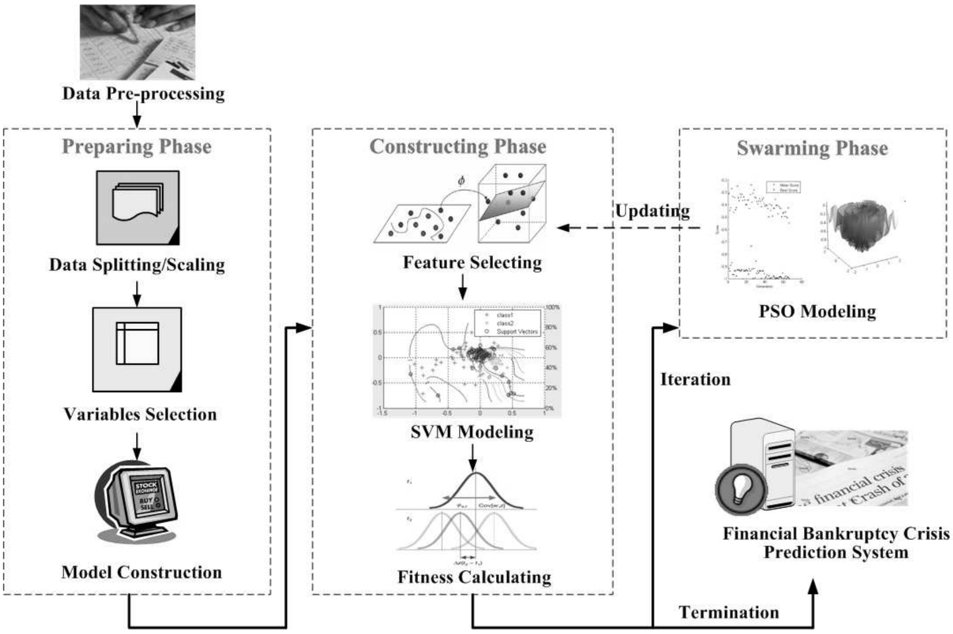

Fig. 4 *Research methodology.*

prediction system are the preparation, construction, and swarming phases. The preparation phase deals with the dataset – basically a huge set of original records from the Taiwan Stock Exchange Corporation (TSEC) which must first undergo data pre-processing. Various methods can be used for preprocessing. In the *splitting* method, we adopted k-fold cross validation and set k as 3, 4, and 5. The data was divided into three, four, and five portions, respectively. In the *transformation* method, the raw data is manipulated to produce a single input. In our experiments, the financial ratios and business cycle index are all number variables which do not require transformation. The *scaling* method applies to prevent feature values in greater numeric ranges from dominating those in smaller numeric ranges. We used the min-max data scaling method to standardize all ratios as Z-values. Moreover, the Z-values were beyond the range of [0, 1], allowing us to prevent numerical difficulties in the calculation. The *denoising* method removes noise from the data. The goal of this phase was to select suitable indicators, including financial ratios and business cycle index, and then construct two models: *Financial Model* and *Business Cycle Model*. In the Financial Model, these input variables were composed of pure financial ratios. However, the financial ratios and business cycle index were collected together in the Business Cycle Model. Once the above processes are completed, the next phase loads the suitable indicators and the discovery prediction patterns for use in the SVM classification.

In the construction phase, we collected financial statement datasets and business cycle index which were derived from the above preparation phase. During the feature selection process, each particle denotes a selected subset of features and parameter values. The selected features, parameter values, and training dataset are used to build the SVM model. After the classification model is built, the testing dataset is used to determine its fitness value, with a higher classification accuracy rate indicating a higher fitness value for the particle. If the particle's fitness is better than its best previous experience (i.e. *pbest)*, the best previous experience of the particle is updated accordingly. Furthermore, if the particle's fitness is better than the global best fitness (i.e. *gbest*), the global best fitness is also updated. If the termination criteria are satisfied, then the process ends, producing the accuracy rate from our proposed financial bankruptcy crisis prediction system; otherwise, the next iteration occurs and the system enters into a swarming phase. The termination criterion used in this research is the maximum number of iterations.

In the swarming phase, the system first initializes the particles and sets the PSO parameters including the feature mask, C, and γ. The PSO parameters set includes the number of iterations, velocity limitation, number of particles, particle dimensions, and weight for fitness calculation. Next, iteration is set to 0, and the training process is performed from "feature selection" through "SVM modeling" to "fitness calculation". Third, the system generates different parameters for the particles and updates the global and personal best values according to the fitness evaluation results. Each particle will move to its next position using formulas (12) and (13). If the stopping criteria or predefined maximum iteration count are met, then the system returns to the construction phase. Finally, with the termination of the training iteration determined in the previous step, the retraining of the PSO will reveal the best values for the SVM parameters feature mask, C, and γ. If the swarming phase is terminated, then the PSO-SVM system will return to the construction phase and use these optimized values obtained from previous PSO training. The SVM model also measures testing accuracy on the testing dataset via the trained SVM classifier, and ends the procedure.

3.2 Data

Our samples contained raw data from 80 electronic firms listed by the Taiwan Stock Exchange Corporation (TSEC). The sampling period ranged from January 2000 to June 2010 (10 years and 6 months). 40 electronic firms in financial distress were matched with 40 healthy electronic firms identified by the absence of any indication or proof concerning issues of financial distress in auditors' reports. All the financial variables used in the sample were extracted from formal financial statements, such as balance sheets, cash flow statements, and income statements. In addition, the business cycle information used in this research was obtained from government annual reports. Thus the utility of this research is not restricted by the limitation of the sample to Taiwanese companies.

3.3 Variables

The selection of variables to be used as candidates for participation in the input vector was based upon prior research linked to predictions of financial distress. From a financial perspective, the selection was based on research by Chen and Du [18] which contains suggested indicators of financial distress prediction. Therefore, this paper adopted related variables based on prior research, the Taiwanese Economic Journal (TEJ), and the Taiwanese Economic Database. The 13 variables selected were Debt to Equity Ratio, Gearing Ratio, Debt/Equity (DE), Return on Asset (ROA), Earnings per Share (EPS), Return on Equity (ROE), Current Ratio, Acid-Test Ratio, Current Assets to Total Assets, Cash Flow to Total Debt Ratio, Cash Flow Ratio, Inventory to Total Assets Ratio, and Inventory to Sales Ratio. Therefore, these 13 variables were collected to set up the *Financial Model* in our experiment.

From a macroeconomic perspective, we extended the financial ratios to include the macroeconomic indexes for financial crisis prediction. These 9 variables include Monetary Aggregates M1B, Direct and Indirect Finance, Stock Price Index, Industrial Production Index, Nonagricultural Employment, Customers-Cleared Exports, Imports of Machinery and Electrical Equipment, Manufacturing Sales, Wholesale, Retail and Food Services Sales. The above variables are expressed in terms of percentage changes over 1-year time span. All variables, except the stock price index, were seasonally adjusted. In this experiment, these variables were normalized and the 9 indexes were recalculated into only 1 "Business Cycle Index". The macroeconomic indexes were obtained from the Council for Economic Planning and Development in Taiwan. Therefore, above 13 financial variables and business cycle index were collected together and set up the Business Cycle Model in our experiment.

4. Empirical Analysis

4.1 Experimental environment and parameter settings

The platform adopted to develop the proposed PSO–SVM and other AI approaches uses a PC with the following features: Intel I7 six-core CPU, 4G RAM, a Windows 7 operating system and the MATLAB R2009 development environment. Throughout the initial experiment, the parameter values used in the proposed PSO–SVM with feature selection were set as follows. The cognition learning factor $c1$ and the social learning factor $c2$ for PSO–SVM were set to 2 and 1, respectively. The number of particles and the maximum number of iterations were set as 20 and 1000, respectively. The searching range of SVM parameter C was set between 1 and 100, while the searching range of the SVM parameter γ was set between 1 and 100 [31]. The k-fold approach [32] was used to evaluate the classification accuracy rate. This research set k as 3, 4, and 5 and the data was divided into three, four, and five portions, respectively. The final average accuracy rate was the average of the three, four, and five accuracy rates.

This process uses the financial and macroeconomic ratios to construct a financial distress prediction model following feature selection strategy. The variables are then loaded as SVM input nodes. In addition, to ensure stability and fairness in the prediction accuracy, we also applied the above experimental parameters to investigate the two, four, six and eight quarters prior to the onset of financial distress.

4.2 Comparative research

4.2.1 Experiments without feature selection strategy

The results obtained by the proposed PSO–SVM approach for Financial Model are shown in Tab. I. To verify the proposed PSO–SVM approach, the results compared by Grid-SVM and GA-SVM are also shown in Tab. I. Besides, this experiment obtained a result with 13 financial ratios. As shown in Tab. I, for the previous two quarters the PSO-SVM produced an estimated average accuracy rate for CV as high as 78.12%. Unexpectedly, the CV average accuracy rates rise to 91.50%, when measured over the previous 8 quarters. As shown in Tab. I, overall durations (two, four, six and eight quarters) PSO-SVM results without feature selection were superior to those of the Grid-SVM and GA-SVM. The average accuracy rates for PSO-SVM, Grid-SVM, and GA-SVM are 87.02%, 86.78%, and 86.65%, respectively. The Fig. 5 shows the PSO-SVM accuracy rate snapshot for two quarters with 3-fold simulation. Besides, the Fig. 6 and Fig. 7 show the Grid-SVM and GA-SVM accuracy rate snapshot for two quarters with simulation, respectively.

This experiment obtained a result after using 13 original pairs of financial ratio variables and business cycle index that had undergone without feature selection in the Business Cycle Model in Tab. II. As shown in Tab. II, for the previous two quarters, the PSO-SVM, Grid-SVM, and GA-SVM have estimated average accuracy rates for CV as high as 89.37, 89.79%, and 87.91%, respectively. However, the CV average accurate rates drop to 86.14%, 86.40%, and 84.73%, respectively, when measured over the previous eight seasons. Tab. II shows the average cross validation with the 3-fold, 4-fold, and 5-fold results for PSO-SVM is better than that produced by GA-SVM. The average accuracy rates for PSO-SVM are 89.37%, 86.97%, 85.55% and 86.14%, respectively, as compared to 87.91%, 86.45%, 83.95% and 84.73% for GA-SVM. Moreover, the evaluation performance between PSO-SVM and Grid-SVM is very similar and has not significant difference.

4.2.2 Experiments with feature selection strategy

This experiment obtained a result after using 13 original pairs of financial ratio variables that had undergone feature selection. As shown in Tab. III, for the previous two quarters in the Financial Model, the CV produced estimated average accuracy rates for PSO-SVM as high as 82.70%. However, the CV average accurate rates drop to 78.00%, when measured over the previous eight quarters and, accuracy for PSO-SVM, Grid-SVM and GA-SVM rises as the time to the financial crisis falls. The average accuracy rates for PSO-SVM, Grid-SVM, and GA-SVM are 81.10%, 80.40%, and 79.34%, respectively. As the result, the PSO-SVM produces a higher CV average accuracy rate than GA-SVM and nearly equal to that of Grid-SVM.

This experiment obtained a result after using 13 original pairs of financial ratio variables and business cycle index that had undergone with feature selection in the Business Cycle Model in Tab. IV. As shown in Tab. IV, for the previous two quarters, the PSO-SVM, Grid-SVM, and GA-SVM have estimated average accuracy rates for CV as high as 89.79, 89.58%, and 88.95%, respectively. However, the CV average accurate rates drop to 73.17%, 71.92%, and 71.45%, respectively, when measured over the previous eight seasons and, in keeping with the results from the

Financial Model	Fold	PSO-SVM			Grid-SVM			GA-SVM		
		c	g	Accuracy	c	g	Accuracy	c	g	Accuracy
2 Quarter CV	3	22.22	0.94	75.62	16	0.70	75	22.01	0.93	75.62
	4	9.72	1.61	78.75	11.31	2	77.5	20.76	1.63	78.75
	5	13.44	1.55	80	181.01	0.70	79.37	3.48	4.10	78.12
	Avg.			**78.12**			**77.29**			**77.50**
4 Quarter CV	3	61.95	0.58	88.12	64	0.5	87.81	52.45	0.60	88.12
	4	71.49	0.94	90.31	64	1	90.31	44.92	0.81	90.31
	5	31.70	1.88	89.06	45.25	0.70	88.75	44.88	0.76	88.75
	Avg.			**89.16**			**88.95**			**89.06**
6 Quarter CV	3	87.38	1.35	88.12	256	1	88.54	20.19	1.58	89.16
	4	5.12	1.58	89.16	128	1.41	89.58	3.12	2.92	87.29
	5	9.08	1.23	90.62	8	1.41	89.79	60.16	2.23	89.37
	Avg.			**89.30**			**89.30**			**88.60**
8 Quarter CV	3	39.14	1.37	90.31	256	0.5	90.46	77.01	1.04	90.46
	4	78.55	1.18	92.03	22.62	1	92.34	9.42	1.57	91.87
	5	13.13	1.86	92.18	22.62	1.41	92.03	26.76	1.19	92.03
	Avg.			**91.50**			**91.61**			**91.45**
Average				**87.02**			**86.78**			**86.65**

Tab. I *Comparison between the PSO-SVM, Grid-SVM, GA-SVM for Financial Model without Feature Selection.*

Business Cycle Model	**Fold**	**PSO-SVM**			**Grid-SVM**			**GA-SVM**		
		c	**g**	**Accuracy**	**c**	**g**	**Accuracy**	**c**	**g**	**Accuracy**
2 Quarter CV	3	2.50	0.53	89.37	2.82	0.5	89.37	65.23	0.40	89.37
	4	2.36	0.58	89.37	1.41	0.70	89.37	54.77	1.62	85
	5	4.31	0.70	89.37	5.65	0.50	90.62	82.02	0.03	89.37
	Avg.			**89.37**			**89.79**			**87.91**
4 Quarter CV	3	79.20	0.21	86.87	256	0.12	87.5	79.32	0.23	86.56
	4	100	0.18	86.56	128	0.17	86.87	74.43	0.22	86.25
	5	52.52	0.21	87.50	45.25	0.25	87.18	76.32	0.34	86.56
	Avg.			**86.97**			**87.18**			**86.45**
6 Quarter CV	3	100	0.19	83.95	256	0.12	85	98.47	0.46	81.45
	4	100	0.53	86.87	181.01	0.35	87.29	47.96	0.68	85.83
	5	41.47	0.23	85.83	181.01	0.35	86.45	88.07	0.53	84.58
	Avg.			**85.55**			**86.24**			**83.95**
8 Quarter CV	3	83.78	0.19	84.21	181.01	0.12	84.21	56.73	0.09	82.96
	4	79.36	0.35	87.34	256	0.17	87.34	11.11	0.96	84.84
	5	74.37	0.24	86.87	256	0.17	87.65	45.39	0.30	86.40
	Avg.			**86.14**			**86.40**			**84.73**
Average				**87.00**			**87.40**			**85.76**

Tab. II *Comparison between the PSO-SVM, Grid-SVM, GA-SVM for Business Cycle Model without Feature Selection.*

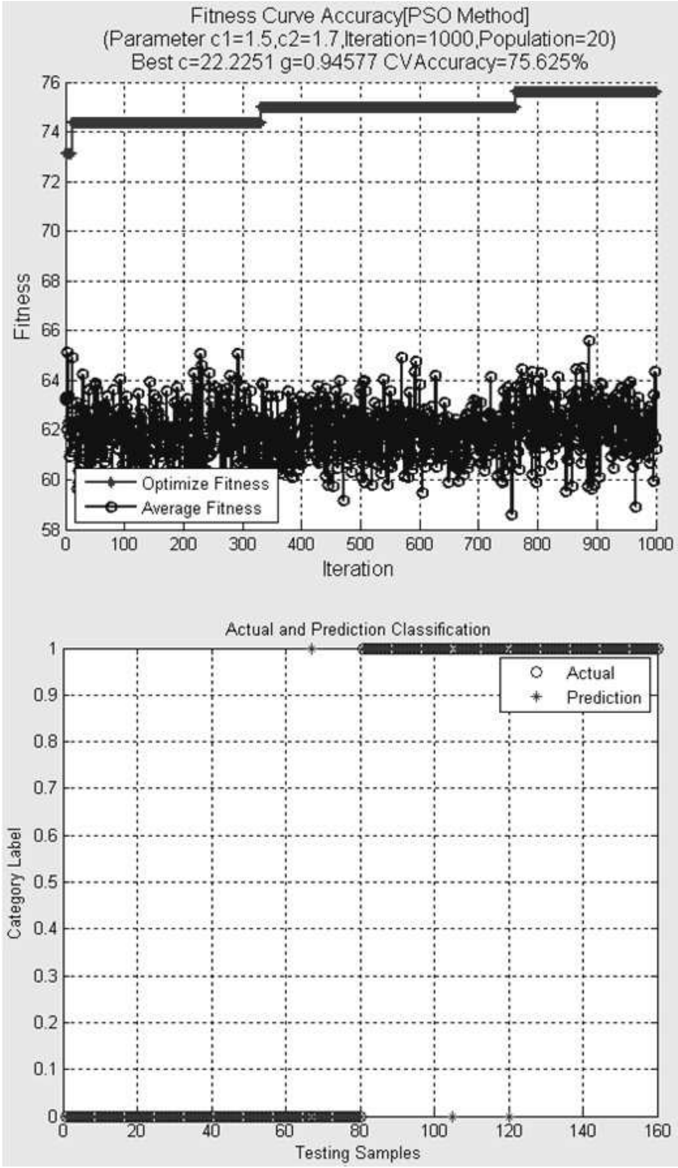

Fig. 5 *The snapshot for PSO-SVM simulation.*

above experiment, accuracy rises as time to the financial crisis falls. Tab. IV shows the average cross validation with the 3-fold, 4-fold, and 5-fold results for PSO-SVM is better than that produced by Grid-SVM and GA-SVM. The average accuracy rates for PSO-SVM, Grid-SVM, and GA-SVM are 80.26%, 79.43%, and 78.77%, respectively.

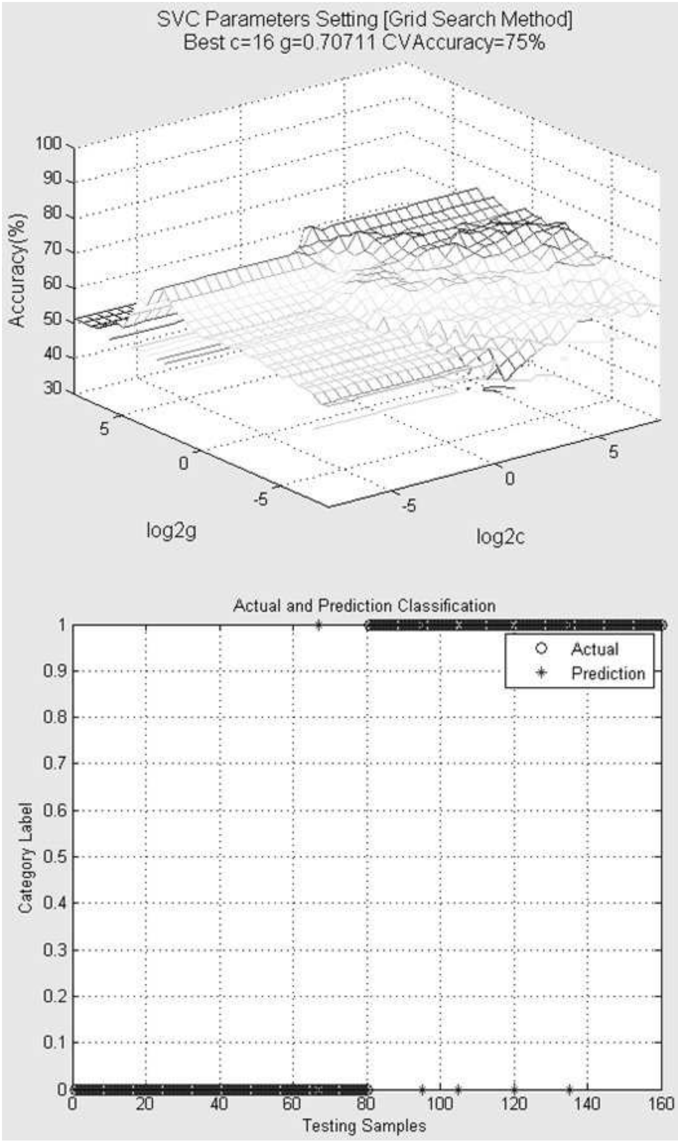

Fig. 6 *The snapshot for Grid-SVM simulation.*

4.3 Discussion

Firstly, we observed the average accuracy rate has no significant difference for Financial Model and Business Cycle model without feature selection strategy in Tab. I and Tab. II. In Tab. I, the average accuracy rates for PSO-SVM, Grid-SVM, and GA-SVM are 87.02%, 86.78%, and 86.65%, respectively. Moreover, the average accuracy rates for PSO-SVM, Grid-SVM, and GA-SVM are 87.00%,

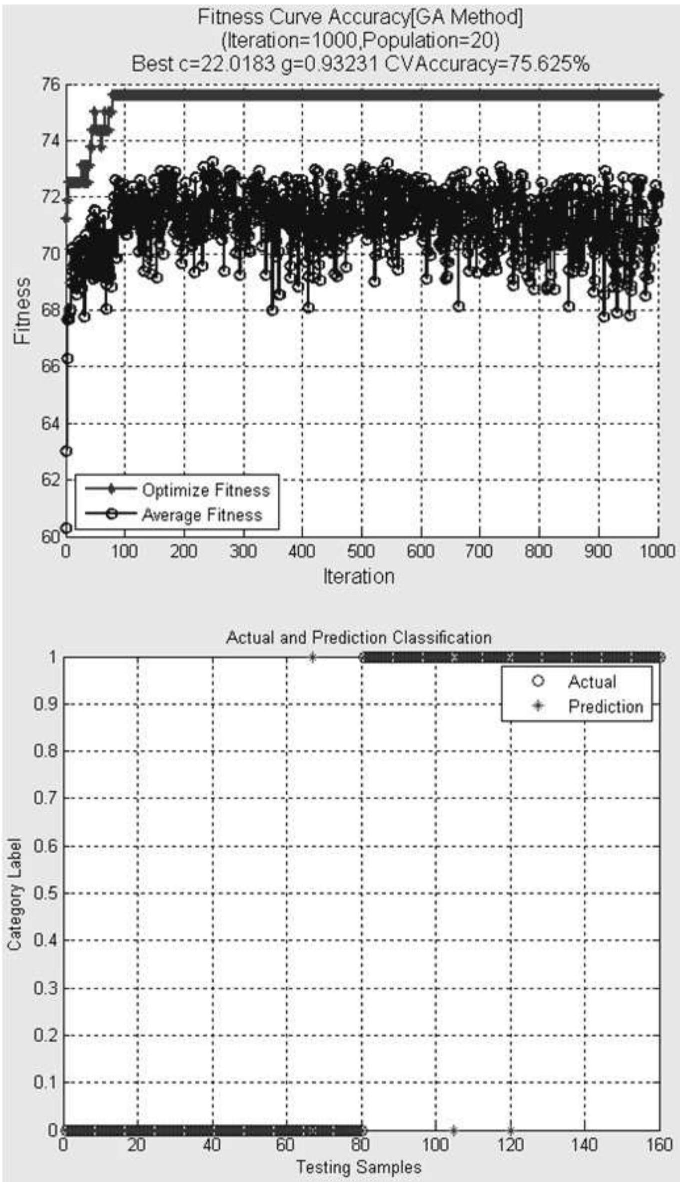

Fig. 7 *The snapshot for GA-SVM simulation.*

87.40%, and 85.76% respectively in Tab. II. However, we could find the PSO-SVM produced an estimated average accuracy rate for CV as high as 78.125% for the previous two quarters in Tab. I. Unexpectedly, the CV average accuracy rates rise to 91.50%, when measured over the previous 8 quarters. The same results obtained as Grid-SVM and GA-SVM in Tab. I. Therefore, we could infer that better accuracy rate would be obtained for PSO-SVM, Grid-SVM and GA-SVM with more datasets when without feature selection. However, we could also observe the different results in Business Cycle Model in Tab. II. As shown in Tab. II, for the previous two quarters, the PSO-SVM, Grid-SVM, and GA-SVM have estimated average accuracy rates for CV as high as 89.37, 89.79%, and 87.91%, respectively. These results have obviously better performance than Financial Model in Tab. I. Therefore, we could also infer that the business cycle index would affect heavy impact for short-term financial bankruptcy prediction.

Secondly, the classification accuracy for PSO-SVM, Grid-SVM, and GA-SVM approaches could evidently be retained by removing noisy and keeping highly correlated features. However, the higher deduction by feature selection results in lower CV classification accuracy for PSO-SVM, Grid-SVM, and GA-SVM. As shown in Tab. III, the average CV accuracy rate for PSO-SVM, Grid-SVM, and GA-SVM drops to 81.10%, 80.40% and 79.34%, respectively, while the average CV accuracy rate for these three approaches is 87.02%, 86.78% and 86.65% in Tab. I, respectively. The same situation results for PSO-SVM, Grid-SVM, and GA-SVM, with average CV accuracy rates 80.26%, 79.43%, and 78.77%, respectively in Tab. IV, while the average CV accuracy rate for these three approaches is 87.00%, 87.40% and 85.76% in Tab. II, respectively. Therefore, the results show that we cannot obtain higher classification accuracy by removing too many variables with feature selection. In our experiments, we could observe that the original 13 features are decreased to 4 features over the previous two quarters in Tab. III. Besides, the original 14 features are also decreased to 4 features over the previous two quarters in Tab. IV. On the contrary, doing so would result in the loss of important or suitable input variables for learning and training with neural networks or soft computing approaches. The only exception is the use of feature selection to extract useful features for short-term financial predictions. Table IV indicates that the PSO-SVM, Grid-SVM, and GA-SVM average accuracy rate for the previous two quarters is not affected by the feature selection. The average accuracy rates for the previous two quarters with PSO-SVM, Grid-SVM, and GA-SVM are 89.79%, 89.58%, and 88.95%, respectively. Therefore, we observe that we could still obtain high average accuracy rates with less features and computation times by using feature selection in Business Cycle Model.

This research presents several key findings regarding the implications and determinants of financial predictions of business bankruptcy:

1. Our approach requires 70% fewer financial ratios than other methods but still presents highly-accurate short-term financial bankruptcy predictions.

2. The experimental results of the feature selection show that our proposed swarm-inspired optimization approach has a high average accuracy rate both for the previous two and four quarters. Thus, the experiments indicate that we could still maintain an acceptable average accuracy rate for long-term predictions, including an accuracy rate above 70% accuracy rate two years (eight quarters) before bankruptcy occurs. Retaining enough suitable features will allow for a higher accuracy rate in the feature selection strategy.

3. This research found that the business cycle index would significantly influence the prediction accuracy, especially in short term. This issue has been seldom explored in prior researches on business failure prediction.

Financial Model	Fold	PSO-SVM				Grid-SVM				GA-SVM			
		c	g	Features	Accuracy	c	g	Features	Accuracy	c	g	Features	Accuracy
2 Quarter CV	3	5.21	1.71	13 → 4	83.75	8	1.41	13 → 4	83.75	4.73	29.50	13 → 4	80.62
	4	4.86	2.13	13 → 4	81.87	5.65	2	13 → 4	81.87	3.74	2.16	13 → 4	80.62
	5	6.59	2.14	13 → 4	82.50	11.31	1.41	13 → 4	82.50	11.25	1.70	13 → 4	83.12
	Avg.				**82.70**				**82.70**				**81.45**
4 Quarter CV	3	13.61	1.10	13 → 4	82.75	11.31	0.70	13 → 4	82.43	6.75	2.40	13 → 4	81.12
	4	8.13	1.36	13 → 4	82.75	8	1.41	13 → 4	82.43	1.85	24.45	13 → 4	80.43
	5	11.04	1.38	13 → 4	82.12	22.62	0.70	13 → 4	82.12	6.69	6.37	13 → 4	80.25
	Avg.				**82.54**				**82.32**				**80.60**
6 Quarter CV	3	6.26	1.41	13 → 4	80.41	45.25	1	13 → 4	79.68	42.90	1.07	13 → 4	78.75
	4	7.18	1.39	13 → 4	81.68	22.62	11.31	13 → 4	79.09	39.55	8.62	13 → 4	78.08
	5	9.65	2.42	13 → 4	81.25	2	22.6	13 → 4	78.74	2.08	23.35	13 → 4	78.12
	Avg.				**81.13**				**79.44**				**78.31**
8 Quarter CV	3	16.54	1.26	13 → 4	77.68	181	0.25	13 → 4	77.68	6.49	2.38	13 → 4	76.37
	4	5.73	1.95	13 → 4	78.31	5.65	2	13 → 4	77.15	2.10	6.54	13 → 4	76.21
	5	47.64	3.89	13 → 4	78.00	32	1	13 → 4	76.84	2.32	13.64	13 → 4	75.50
	Avg.				**78.00**				**77.22**				**76.02**
Average					**81.10**				**80.40**				**79.34**

Tab. III *Comparison between the PSO-SVM, Grid-SVM, GA-SVM for Financial Model with Feature Selection.*

Business Cycle Model	Fold	PSO-SVM				Grid-SVM				GA-SVM			
		c	**g**	**Features**	**Accuracy**	**c**	**g**	**Features**	**Accuracy**	**c**	**g**	**Features**	**Accuracy**
2 Quarter CV	3	64.59	0.46	14 → 4	90.62	181.1	0.25	14 → 4	90.00	68.66	0.39	14 → 4	90.00
	4	25.68	0.13	14 → 4	89.37	90.50	0.12	14 → 4	89.37	11.89	4.17	14 → 4	88.75
	5	13.55	5.13	14 → 4	89.37	16	0.12	14 → 4	89.37	3.6	10.19	14 → 4	88.12
	Avg.				**89.79**				**89.58**				**88.95**
4 Quarter CV	3	2.42	0.92	14 → 4	84.62	2.82	1	14 → 4	82.93	4.42	0.61	14 → 4	82.31
	4	28.03	1.26	14 → 4	83.81	4	0.70	14 → 4	82.93	3.95	0.66	14 → 4	81.93
	5	12.16	0.42	14 → 4	83.75	4	1	14 → 4	82.31	29.06	0.95	14 → 4	81.68
	Avg.				**84.06**				**82.72**				**81.97**
6 Quarter CV	3	13.95	0.40	14 → 4	73.66	4	1	14 → 4	73.50	12.07	0.52	14 → 4	72.66
	4	51.62	0.01	14 → 4	74.37	4	1	14 → 4	73.70	3.02	1.56	14 → 4	72.95
	5	7.25	0.86	14 → 4	74.04	2	1.41	14 → 4	73.33	4.47	1.12	14 → 4	72.54
	Avg.				74.02				**73.51**				72.71
8 Quarter CV	3	24.18	0.91	14 → 4	72.65	32	0.25	14 → 4	71.15	4.52	0.63	14 → 4	71.00
	4	15.43	1.33	14 → 4	73.43	0.70	2	14 → 4	72.37	77.99	0.52	14 → 4	71.84
	5	5.34	1.39	14 → 4	73.43	22.62	0.35	14 → 4	72.25	17.48	0.82	14 → 4	71.53
	Avg.				**73.17**				**71.92**				**71.45**
Average					**80.26**				**79.43**				**78.77**

Tab. IV *Comparison between the PSO-SVM, Grid-SVM, GA-SVM for Business Cycle Model with Feature Selection.*

4. This study empirically determined that the swarm-inspired optimization achieved better forecasting accuracy than other evolutionary approaches, such as Grid search and GA. Furthermore, the swarm-inspired optimization approach has rarely been used for forecasting, especially for financial problems.

Thus, the findings of this research are valuable and provide several important implications for research in financial predictions and in practice.

4.4 UCI datasets

To compare the proposed PSO–SVM with other approaches, we used three financial related datasets, two popular small datasets, and two large datasets collected from the UCI Machine Learning Repository [33]. The seven datasets are Australian Credit Approval, German Credit Data, Japan CRX, Iris, Wine, Chess, and Adult, respectively. The other approaches used as a basis for comparison were SVM- and SOM-related approaches which reported at least one classification accuracy rate for the above-mentioned datasets in the literature. These approaches included Grid-SVM, GA-SVM, SVM, and SOM. For the purposes of discussion, the experimental parameters settings were identical to those detailed in Section 4.1, and the feature selection methods for all related-SVM approaches were adopted and run to measure their relative prediction performance. Tab. V clearly shows that the proposed PSO–SVM provides the highest accuracy rate for the seven UCI datasets. The experimental results show that there are no significant differences between PSO-SVM and Grid-SVM when handling small datasets with few variables. However, PSO-SVM has better prediction rate than Grid-SVM for datasets with larger instances and more variables, as shown in Tab. V. This demonstrates that the proposed approaches can be applied to financial and other widely-used datasets.

Dataset (From UCI)	Variables	Instances	PSO--SVM	Grid--SVM	GA--SVM	SVM	SOM
Japanese	15	125	86.16%	86.16%	82.78%	81.63%	65.82%
Iris	4	150	97.77%	97.77%	95.33%	96.00%	91.33%
Wine	13	178	100%	100%	91.66%	89.88%	97.75%
Australian	14	690	85.10%	86.06%	84.62%	85.02%	73.43%
German	20	1000	80.33%	79%	78.33%	75.33%	73.33%
Chess	36	3196	90.11%	83.54%	81.20%	80.03%	72.15%
Adult	14	48842	92.00%	86.01%	84.99%	82.00%	76.19%

Tab. V *Comparison UCI datasets between various approaches.*

5. Conclusions

This research focused on 13 financial ratios and 1 business cycle index used in financial statements, and used the PSO-SVM, Grid-SVM and GA-SVM to compare the performance of financial distress predictions. 40 electronic companies in financial crisis were matched with 40 healthy electronic companies in the electronic industries. The dataset was obtained from the TSEC database and sampled them for durations of two, four, six and eight quarters prior to the onset of financial crisis. This data was then used to establish Financial Model and Business Cycle Model, with comparisons made for each ratio variable in the models.

Our experiments provide four critical contributions. Firstly, we found that when we applied business cycle index to construct Business Cycle Model, we could obtain better accuracy rate than Financial Model in short-term financial bankruptcy prediction. This result shows the macroeconomic index mainly affects financial prediction performance and has greater influence than financial ratios.

Secondly, the closer we get to the onset of actual financial distress, the more accurate the prediction will be in the PSO-SVM, Grid-SVM, and GA-SVM approaches with business cycle index. For example, PSO-SVM without feature selection provides 89.37% average CV accuracy rate for two quarters prior to the occurrence of financial distress, but only 86.14% when measured eight quarters in advance. However, the larger dataset we own the higher accuracy rate we get, when we do not consider adopting business cycle index and featuring selection.

Thirdly, the PSO-SVM approach yields higher classification accuracy than other approaches. Removing those noisy and highly-correlated features greatly improves computation times, but removing too many features would adversely affect classification accuracy. According to the literature, when adopting a crisis prediction model, most researchers extracted and sorted reduction variables for convenient analysis. However, the experimental results show that, given a small number of variables, an error exists between the reduction and original variables, and that artificial intelligence or neural networks would not have enough data to conduct learning and training. Therefore, the experiments show that we could still retain high average accuracy rate if we keep enough and suitable features.

Finally, the PSO-SVM approach generally produces better prediction accuracy than the Grid-SVM, GA-SVM, SVM, SOM, and SVR-SOM models in predicting the Australian, German, Japan CRX financial datasets, Iris, Wine, Adult, and Chess datasets from UCI. Therefore, our proposed PSO-SVM approach is suitable for predicting financial distress and events in other fields.

More research is needed. Firstly, while the results in this research were obtained through the PSO method, other soft-computing methods can also be applied to the SVM-based approach. Secondly, the experimental results obtained from other public datasets or real-world problems can be tested to verify and extend this approach.

Acknowledgements

The author thanks the support of the National Scientific Council (NSC) of the Republic of China (ROC) to this work under Grant No. NSC-98-2410-H-025-011. The author also gratefully acknowledges the Editor and anonymous reviewers for their valuable comments and constructive suggestions.

References

[1] Garay U.: The Asian Financial Crisis of 1997 – 1998 and the Behavior of Asian Stock Markets. A Web Journal of Applied Topics in Business and Economics, 2003. Accessed available from http://www.westga.edu/~bquest/2003/asian.htm

[2] Chen M. H.: How Could Taiwan Have been Insulated from the 1997 Financial Crisis? National Policy Foundation Research Report, National Policy Foundation, 2001.

[3] Business Wire, Three top economists agree 2009 worst financial crisis since great depression; risks increase if right steps are not taken. February, 29, 2009. Business Wire News database. Accessed 16 August 2010.

[4] Jaffe M.: Government Watchdog Says Treasury and Fed Knew Bailed-Out Banks Were Not Healthy. ABC News Dataset, October 5, 2009.

[5] FDIC, Failed Bank List, 2010. Available from the website database: http://www.fdic.gov/bank/individual/failed/banklist.html

[6] Min J. H., Jeong C.: A binary classification method for bankruptcy prediction. Expert Systems with Applications **36**, 2009, pp. 5256–5263.

[7] Beaver W.: Financial ratios as predictors of failure, empirical research in accounting: Selected studies. Journal of Accounting Research, 1966, pp. 71–111.

[8] Altman E. L.: Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance, **23**, 3, 1968, pp. 589–609.

[9] Martin D.: Early warning of bank failure a logit (?) regression approach. Journal of Banking & Finance, **1**, 1977, pp. 249–276.

[10] Zmijewski M. E.: Methodological issues related to the estimation of financial distress prediction models. Journal of Accounting Research, **22**, 1984, pp. 59–82.

[11] Cielen A., Peeters L., Vanhoof K: Bankruptcy prediction using a data envelopment analysis. European Journal of Operational Research, **154**, 2004, pp. 526–532.

[12] Kao C., Liu S. T.: Prediction bank performance with financial forecasts: A case of Taiwan commercial banks. Journal of Banking & Finance, **28**, 2004, pp. 2353–2368.

[13] Hu Y. C., Tseng F. M.: Functional-link net with fuzzy integral for bankruptcy prediction. Neurocomputing, **70**, 16–18, 2007, pp. 2959–2968.

[14] Sanchis A., Segovia M. J., Gil J. A., Heras A., Vilar J. L.: Rough Sets and the role of the monetary policy in financial stability and the prediction of insolvency in insurance sector. European Journal of Operational Research, **181**, 3, 2007, pp. 1554–1573.

[15] Kirkos E., Spathis C., Manolopoulos Y.: Data mining techniques for the detection of fraudulent financial statements. Expert Systems with Applications, **32**, 4, 2007, pp. 995–1003.

[16] Sun J., Li H.: Data mining method for listed companies' financial distress prediction. Knowledge-Based Systems, **21**, 1, 2008, pp. 1–5.

[17] Ahn H., Kim K.: Bankruptcy prediction modeling with hybrid case-based reasoning and genetic algorithms approach. Applied Soft Computing, **9**, 2, 2009, pp. 599–607.

[18] Chen M. Y., Du Y. K.: Using neural networks and data mining techniques for the financial distress prediction model. Expert Systems with Applications, **36**, 2, 2009, pp. 4075–4086.

[19] Boyacioglu M. A., Kara Y., Baykan Ö. K.: Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: A comparative analysis in the sample of savings deposit insurance fund (SDIF) transferred banks in Turkey. Expert Systems with Applications, **36**, 2, 2009, pp. 3355–3366.

[20] Jang J. S. R., Sun C. T., Mizutani E.: Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Prentice-Hall, New Jersey, 1997.

[21] Panda S. S., Chakraborty D., Pal S. K.: Flank wear prediction in drilling using back-propagation neural network and radial basis function network. Applied Soft Computing, **8**, 2, 2008, pp. 858–871.

[22] Vapnik V. N.: Statistical Learning Theory. Wiley. New York, 1998.

[23] Drucker H., Burges J. C., Kaufman L., Smola A., Vapnik V.: Support Vector Regression Machines. Advances in Neural Information Processing Systems, **9**, 1966, pp. 155–161.

[24] Hsu S. H., Hsieh P. A., Chih T. C., Hsu K. C.: A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression. Expert Systems with Applications, **36**, 2009, pp. 7947–7951.

[25] Eberhart R. C., Kennedy J.: A new optimizer using particle swarm theory. Proc. Sixth International Symposium on Micro Machine and Human Science (Nagoya, Japan). IEEE Service Center, Piscataway, NJ, 1995, pp. 39–43.

[26] Yuan S. F., Chu F. L.: Fault diagnostics based on particle swarm optimization and support vector machines,. Mechanical Systems and Signal Processing, **21**, 4, 2007, pp. 1787–1798.

[27] Kennedy J., Eberhart R.: Particle swarm optimization. Proc. IEEE International Conf. on Neural Networks (Perth, Australia). IEEE Service Center, Piscataway, NJ, 1995.

[28] Shi Y., Eberhart R. C.: A modified particle swarm optimizer. In: Proceedings of the IEEE International Conference on Evolutionary Computation, Piscataway, NJ, 1998, pp. 69–73.

[29] Angeline P. J.: Using selection to improve particle swarm optimization. IEEE International Conference on Evolutionary Computation, Anchorage, Alaska, May 1998, pp. 4–9.

[30] Eberhart R. C., Shi Y. H.: Comparison between genetic algorithms and particle swarm optimization. 1988 Annual Conference on Evolutionary Programming, San Diego, 1998.

[31] Chang C. C., Lin C. J.: LIBSVM: a library for support vector machines, 2001. Software available at `http://www.csie.ntu.edu.tw/∼cjlin/libsvm`, Accessed 16 August 2010.

[32] Salzberg S. L.: On comparing classifiers: Pitfalls to avoid and a recommended approach. Data Mining and Knowledge Discovery, **1**, 1997, pp. 317–327.

[33] Hettich S.: UCI machine learning repository. School of Information and Computer Science, University of California, Irvine, 2004.
`http://www.ics.uci.edu/∼mlearn/MLRepository.html`