A RELEVANCE VECTOR MACHINE WITH ROUGH SET THEORY MODEL IN ANALYZING THE LIFE CYCLE OF NEW ECONOMIC FIRMS

Ming-Fu Hsu, *Ping-Feng Pai,† Wei-Shih Chung†*

Abstract: The subprime mortgage crisis and subsequent financial tsunami have raised considerable concerns about financial risk management and evaluation. This is nowhere more apparent than in new economic firms (NEFs) with large economic targets and heavy R&D expenses, such as firms in the electronics industries. With its potential for extreme growth and superior profitability, the electronic industries in Taiwan have been in the financial stock market spotlight. Recently, the relevance vector machine (RVM) was reported to have considerably less computation complexity than support vector machines (SVM) models, since it uses fewer kernel functions. Another emerging technique is rough set theory (RST), which derives rules from data. Based on the corporation life cycle theory (CLC), this study developed a relevance vector machine with rough set theory (RVMRS) to predict the status of a corporation in the decline stage. To demonstrate the performance of the designed RVMRS model, the study used electronic industries data from the Taiwan Economic Journal data bank, Taiwan Security Exchange, and Securities and Futures Institute in Taiwan. Experimental results revealed that the presented RVMRS model can predict the decline stage in a firm's life cycle with satisfactory accuracy, and generate rules for investors, managers, bankers and regulators that enable them to make suitable judgments. In addition, this study proved that the transparency and information disclosure index (TDI) is crucial to predicting the financial decline of corporations.

*Ming-Fu Hsu
Department of International Business Studies, National Chi Nan University, 1 University Rd., Puli, Nantou 545, Taiwan, ROC, E-mail: s97212903@ncnu.edu.tw
†Ping-Feng Pai – Corresponding Author, Wei-Shih Chung
Department of Information Management, National Chi Nan University, 1 University Rd., Puli, Nantou 545, Taiwan, ROC, Tel.: +886 49-2910960x4672, Fax: +886 49-2915205, E-mail: paipf@ncnu.edu.tw, skamodel@gmail.com

1.   Introduction

In recent years, the sub-prime mortgage crisis and subsequent financial tsunami have weakened the stability of the global financial environment. An accurate assessment of the probability of financial distress is now much more essential and has consequently attracted considerable attention. Due to the high degree of potential growth in the electronic industries, most investors take it into consideration in formulating a personal financial portfolio. According to the definition of Granlund and Taipaleenmäki [1], the electronic industry is one of the new economic firms (NEFs) which include other new or fast-growing firms, such as business in information and communication technology. There exists remarkable dissimilarity between risk assessment and control practices in NEFs and corporations operating in conventional environments. According to Clark and Neill [2], NEFs have been very much in the spotlight recently, as many of them have encountered numerous financial problems.

Most of the extant literature has endeavored to develop effective financial warning models to cope with the financial distress problem; however, few researches have considered the contribution of the corporate life cycle (CLC) theory on this issue. Greiner [3] and Adizes [4] indicated the corporations experience changes throughout their life cycle stages, and the characteristics of corporations quite differ at each phase. The theoretical CLC model is deterministic and indicates that corporations inevitably evolve and transit from one phase of development to another [5]. Thus, the CLC assumes that corporations obey a predictable pattern, which is unable to be reversed easily. After going through the stages of birth, rapid growth and stable maturity, many corporations fall into decline and subsequently encounter distress in many aspects [6]. Black [7] reported that the information about a corporation's earnings and operating cash flow could be expected to differ among corporations in different life cycle phases. Because the decline phase is much more critical than the other life cycle phases, this study aims at analyzing the status of corporations in decline and financial distress.

Pioneer research of financial distress has been done by Beaver [8] and Altman [9], who utilized traditional statistic approaches such as univariate analysis, multiple discriminant analysis and logistic regression to build up an early warning model. Altman [9] and Altman et al. [10] have applied multiple discriminant analysis to cope with the financial distress problem. The logistic regression method is easier to use than the multiple discriminant analysis approach because it requires fewer statistical assumptions and performs a satisfactory job of empirical discrimination [11].

Recently, machine learning techniques have been popularly applied in financial distress tasks with encouraging results. Machine learning is effective in an exploratory analysis scenario where there are no predetermined notions about what will constitute an interesting outcome [12]. Various techniques have been employed including support vector machines [13], rough set theory [14], decision trees [15] and neural networks [16, 17].

The life cycle perspective has rarely been utilized in financial distress studies. Thus, this investigation employed CLC aspects to construct an early warning model, namely the RVMRS model, which can be used to predict and analyze the decline stage of firms. As a Bayesian treatment of the sparse learning task, the relevance vector machine (RVM) [18] possess superior ability of generalization and

yields outcomes in a probabilistic format. RVM can avoid some limitations of kernel-based classifiers. The limitations include the requirement of Mercer kernel and suitable parameters determination of classifiers. In addition, RVM can significantly alleviate the computational complexity by using much fewer kernel functions; and thus the RVM technique has been used to deal with many real-time problems [19, 20, 21]. Due to lack of abilities of interpretation, the interpretability of outcomes is a major drawback of RVM models. Thus, this study employed rough set theory to deal with this problem. The RST [22] is an emerging mathematical way to tackle with vague data. The pre-definition of data is not required when performing RST. The RST approach is able to discover unseen knowledge from data and express the knowledge in easily understandable ways for decision makers.

The rest of this paper proceeds as follows. Section 2 illustrates the methodologies used to construct the proposed model. Section 3 describes a numerical example and the experimental results. Section 4 concludes the paper and provides suggestions for future research.

2. Methodologies

2.1 Sequential forward feature selection

Proposed by Pudil et al. [23], sequential forward feature selection (SFFS) is a typical heuristic searching approach whose main concept is to determine the essential features from unselected features and put them into a selected feature subset iteratively. The SFFS starts at an empty set of features and then determines one feature at a time for each iteration until there is no further improvement in forecasting accuracy. In each iteration, all features without predetermined knowledge are considered for selection, and the selected features are based on an assessment score. The feature with the greatest performance is selected into the subset of candidate features. The K-nearest neighbors (KNN) is specified as a search method; and Eq. (1) expresses the recognition result of the leave-one-out test.

$$Recognition\ rate = (B - b)/B \tag{1}$$

where the number of total training subsets and the number of miss classification are represented as B and b, respectively. The features with superior performance are employed for the next iteration. The computational process is performed until the terminated status is reached. The terminated status is triggered when all the features have been investigated. Suppose that the objective function denoted as F and the feature subsets selected from primordial features set Q at g-th steps are expressed as S_g. The following steps illustrate the SFFS procedure to select the feature q^+.

1. Start at an empty set: $S_0 = \{\phi\}$.

2. Determine the next outperforming feature: $q^+ = \arg\max_{\substack{q \in Q \\ q \in S_g}} \{F(S_g + q)\}$.

3. Revise the procedure: $S_{g+1} = S_g + q^+, g = g + 1$

4. Return to procedure (B) until the terminated status is satisfied.

2.2 Relevance vector machine

Designed by Tipping [18], the relevance vector machine (RVM) is a probabilistic non-linear model based on Bayesian theory that has been proven to provide nearly equivalent performance to support vector machine (SVM) with fewer relevance vectors than support vectors. Therefore, with fewer vectors used in the computing mechanism, RVM models have outperformed SVM models in many investigations [24, 25]. The training sample is set of input-output pairs $(\mathbf{x}_j, y_j)_{j=1}^{G}$, where $\mathbf{x}_j \in \mathbf{R}^g$ represents g-dimensional input variables and $y_j \in \mathbf{R}$ is the corresponding output. Equation (2) is employed to illustrate the outputs sampled from the regression model with error term e_j.

$$y_j = t\left(\mathbf{x}_j; \mathbf{v}\right) + e_j \tag{2}$$

where e_j follows Gaussian distribution, that is $e_j \sim N\left(0, \sigma^2\right)$.

Furthermore, Eq. (3) illustrates the decision function of RVM.

$$t\left(\mathbf{x}; \mathbf{v}\right) = \sum_{j=1}^{G} v_j M\left(\mathbf{x}, \mathbf{x}_j\right) + v_0 \tag{3}$$

where $M\left(\mathbf{x}, \mathbf{x}_j\right)$ indicates a foundation function and $\mathbf{v} = (v_0, \ldots, v_G)^T$ represents weights. In accordance with the noise assumption of e_j and Eq. (2), the y_j with mean $t\left(\mathbf{x}_j; \mathbf{v}\right)$ and the variance σ^2 follows the Gaussian distribution and can be expressed as following Eq. (4).

$$p\left(y_j \left| \mathbf{x}_j\right.\right) = N\left(y_i \left| t\left(\mathbf{x}_i; \mathbf{v}\right), \sigma^2\right.\right) \tag{4}$$

The hyperparameter λ is represented as $\lambda = 1/\sigma^2$. Thus, Eq. (5) illustrates the log-likelihood function of the whole training data.

$$p\left(\mathbf{y} \left| \mathbf{v}, 1/\lambda\right.\right) = \left(\lambda/2\pi\right)^{G/2} \exp\left[-\lambda/2 \left\|\mathbf{y} - \mathbf{\Psi}\mathbf{v}\right\|^2\right] \tag{5}$$

where $\mathbf{y} = (y_1, \ldots, y_G)^T$ and $\Psi \in R^{G \times (G+1)}$ can be rewritten as $\Psi = \{\psi\left(\mathbf{x}_1\right), \ldots, \psi\left(\mathbf{x}_G\right)\}^T$. The distribution over the weight $\mathbf{v}$ is calculated by the training RVM mechanism. The purpose of the prior distribution over v_i is to retain sparsity and maximize the likelihood function. Suppose that v_i follows $N\left(0, \beta_i^{-1}\right)$; Eq. (6) is used to illustrate the prior distribution over $\mathbf{v}$.

$$p\left(\mathbf{v} \left| \alpha\right.\right) = \prod_{i=0}^{G} N\left(v_i \left| 0, \alpha_i^{-1}\right.\right) \tag{6}$$

where α_i is the hyper-parameter related to each weight, v_i. According to Eqs. (5) and (6), the posterior distribution over $\mathbf{v}$ can be obtained by the Bayesian theory:

$$p\left(\mathbf{v} \left| \mathbf{y}, \alpha, \lambda\right.\right) = p\left(\mathbf{v} \left| \alpha\right.\right) p\left(\mathbf{y} \left| \mathbf{v}, \lambda\right.\right) / p\left(\mathbf{y} \left| \alpha, \lambda\right.\right) \tag{7}$$

Both $p\left(\mathbf{v} \left| \beta\right.\right)$ and $p\left(\mathbf{y} \left| \mathbf{v}, \alpha\right.\right)$ are Gaussian distributions, and thus the product of two distributions follows Gaussian distribution. $p\left(\mathbf{y} \left| \beta, \alpha\right.\right)$ is a normalized coefficient which does not include $\mathbf{v}$. Equation (8) is a posterior distribution over $\mathbf{v}$.

$$p\left(\mathbf{v} \left| \mathbf{y}, \alpha, \lambda\right.\right) = N\left(\mathbf{v} \left| \mathbf{z}, \Re\right.\right) \tag{8}$$

where $\mathbf{z} = \alpha \Re \Psi^T \mathbf{y}$ and $\Re = \left(\alpha \Psi^T \Psi + \mathbf{Q}\right)^{-1}$ are used to express the mean vector and variance matrix; and $\mathbf{Q} = diag\left(\beta_0, \ldots, \beta_G\right)$.

The parameters α and β are employed to decide the posterior distribution over $\mathbf{v}$. The optimized equations of parameters are expressed in Eqs. (9) and (10).

$$\beta_i = {}^1\!/\!\left(z_i^2 + \Re_{ii}\right) = {}^{\gamma_i}\!/\!{z_i^2}, i = 0, \ldots, G \tag{9}$$

$$\alpha = \left(G - \Re_{i=0}^G \gamma_i\right)\!/\!{\|\mathbf{y} - \Psi\Re\|} \tag{10}$$

where z_i and $\Re_{ii}$ are used to indicate the i-th element of vector $\mathbf{z}$ and the i-th diagonal element of matrix $\Re$, $\gamma_i = 1 - \beta_i \Re_{ii}$. During the training phase, Eqs. (7)–(10) are conducted until the parameters converge or the pre-set number of iterations is obtained [21].

2.3 Rough set theory

Developed by Pawlak [26], rough set theory (RST) is a helpful mathematical approach that allows decision makers to both identify hidden knowledge in a huge amount of information and generate decision rules. Rough set theory employs information systems to stand for knowledge and deal with vague data; it includes four basic concepts: knowledge expression, set estimation, diminution of knowledge, and rules generation. An information system can be represented by the quintuple $IS = (R, Q, E, f)$, where R consists of a non-empty finite set of objects (namely the universe), Q is a non-empty finite set of attributes to delineate the objects in the universe, E expresses the range of $R \times Q$, and $f : R \times Q \to E$ is an information function where $f(x, a) \in E$ for every $a \in Q$ and $x \in R$ [27]. In addition, it is assumed that $K \subseteq Q$ and $(p, q) \in R^2$.

An essential method in knowledge discovery is indiscernibility, which can be treated as a mathematical tool for acquiring facts from an imperfect dataset. Indiscernibility arises from an inability to differentiate among objects in a distinct set, and it leads to identical information derived from different observations. Indiscernible objects build up a major base of knowledge, which are thus called elementary sets. The symbols p and q are utilized to represent two objects. Indiscernibility means that these two objects cannot be distinguished in a particular set; thus, they are merged as identical information. Eq. (11) represents the indiscernibility relationship of p and q.

$$Indiscernibility\,(K) = \left\{(p, q) \in R^2 \,|\, f(p, a) = f(q, a), \forall a \in K\right\} \tag{11}$$

Objects in elementary sets are those that are clearly classified by condition attributes. Nevertheless, not all objects can be categorized into a set by condition attributes; instead, they can only be approximately defined. Thus, the approximation of a set is used to handle approximate definitions through lower and upper approximations. The approximations are formed by the information system. Lower approximations are objects which definitely belong to a certain subset of interest, whereas upper approximations are objects that do not precisely fit into any of these subsets. Let $K \subseteq Q$ and $P \subseteq R$. Eqs. (12) and (13) express the K-lower approximation of $\underline{P}(K)$ and the K-upper approximation of $\overline{P}(K)$, respectively.

$$\underline{P}(K) = \{p \in R \,|\, Indiscernibility\,(P) \subseteq P\} \tag{12}$$

$$\overline{P}\left(K\right) = \{p \in R \,|\, Indiscernibility\left(P\right) \cap P \neq \phi\} \tag{13}$$

If an object falls into the boundary region of these two approximations, this object can not be categorized into any set. The boundary region of these two approximations is an uncertain region. If an object falls into a boundary region, it is impossible to identify the set to which it belongs. These sets are the indiscernibility relations of elementary sets [$Indiscernibility\left(K\right)$].

In information systems, some attributes are irrelevant or redundant. These attributes can be eliminated without decreasing the approximation quality of an information system with an original set of attributes. Therefore, the indiscernibility relation of attributes is unaffected when redundant attributes are deleted. The reduct is the minimal subset of features assigned a classification equivalent to the performance of the whole set of features. Attributes that do not belong to a reduct are unnecessary in terms of the classification of elements in the universe. The intersection of all reducts generates the core, which is the most important feature in the decision table. Eq. (14) expresses the relation of the reduction and core.

$$Core(Q) = \cap \, Reduction(Q) \tag{14}$$

Decision rules are derived from the information table in an interpretable form and they reveal knowledge. The decision rule represents a relationship between a set of conditions and a conclusion. Mark and Munakata [28] reported that rule extraction by RST is relatively efficient and straightforward. If a certain condition(s) in a decision rule matches, the decision will be conducted. Each decision rule explains only a part of the entire information table, and flexibility in modifying the rules is possible. Each decision rule is specified by a support representing the number of records matching the rule [29]. Eq. (15) illustrates a support of a rule.

$$Support\left(p\right) = Cardinality\left(Q\left(p\right) \cap E\left(p\right)\right) \tag{15}$$

where p denotes the identification of a certain rule, and $Cardinality$ is the cardinality of the set. The rule strength is applied to depict the importance in the information table, and it is expressed as Eq. (16). Decision rules with higher values of strength represent more knowledge or information in the data set.

$$Strength = Support\left(p\right) / Cardinality\left(R\right) \tag{16}$$

According to Pawlak [29] and Wang et al. [30], the coverage measurement is another critical indicator in evaluating the performance of RST. It is depicted by Eq. (17).

$$Coverage(E) = Support(E)/Cardinality(E) \tag{17}$$

where the total amount of objects and the support for a specific decision class are expressed as $Cardinality\left(E\right)$ and $Support\left(E\right)$, respectively. Dimitras et al. [31] stated that accuracy is evaluated by the cardinality ratio between the lower approximation and the upper approximation. Eq. (18) depicts the evaluation of accuracy.

$$Accuracy = \sum Cardinality\left(\underline{P}\left(K\right)\right) / \sum Cardinality\left(\overline{P}\left(K\right)\right) \tag{18}$$

3. Numerical examples and experimental results

3.1 The RVMRS model

The RVMRS model is a financial early warning mechanism for forecasting and analyzing the status of corporations. A detailed framework of the model is depicted in Fig. 1. Following Anthony and Ramesh [32], this study divided the life cycle of a corporation into three stages, growth, maturity and decline, by using sales growth, capital expenditure, the dividend payout ratio, and the firm's age. The sales growth rate and capital expenditure ratios were ranked from the highest to the lowest. The other the two ratios were ranked from the smallest to the largest. Using this ranking procedure, we divided the corporations into three equal-sized groups; values of one, two and three were used as rates for each group based on the evaluation depicted in Tab. I. Thus, corporations with a ranking of 3 were treated as in the decline stage and the others as in non-decline stages.

By using financing policies in advance, mangers and governors can prevent corporations from falling into decline and avert a potential financial crisis. However, conducting this evaluation is not entirely straightforward. A large amount of data may result in inefficient and even inapplicable algorithms [33]. In addition, irrelevant or redundant data can lead to improper conclusions. Thus, this study used the sequential forward feature selection (SFFS) technique to select important attributes of data. After carrying out the feature selection procedure, the study divided the selected data into training data and testing data. The training data were utilized to form a tentative RVM model; and the testing data were used to evaluate the performance of the tentative RVM model by classification accuracy. The study then conducted a 5-fold cross-validation (CV) to moderate the overfitting problem. The advantage of CV is that it minimizes the data dependency and increases the reliability of outcomes [34]. Also, Davis et al. [35] stated that even a limited interpretation can positively influence the model acceptance by end users. Thus, the relevance vectors (RVs) and predicted labels with the best CV result were used to yield rules via the rough set mechanism. Finally, rule pruning was performed to select rules according to their importance, since a condensed rule set is more easily comprehended and used by decision makers.

Status	Stages of life cycle		
	Growth	Mature	Decline
Sales growth rate	High	Medium	Low
Capital expenditure	High	Medium	Low
Dividend payout ratio	Low	Medium	High
Firm's age	Low	Medium	High
Ranking	1	2	3

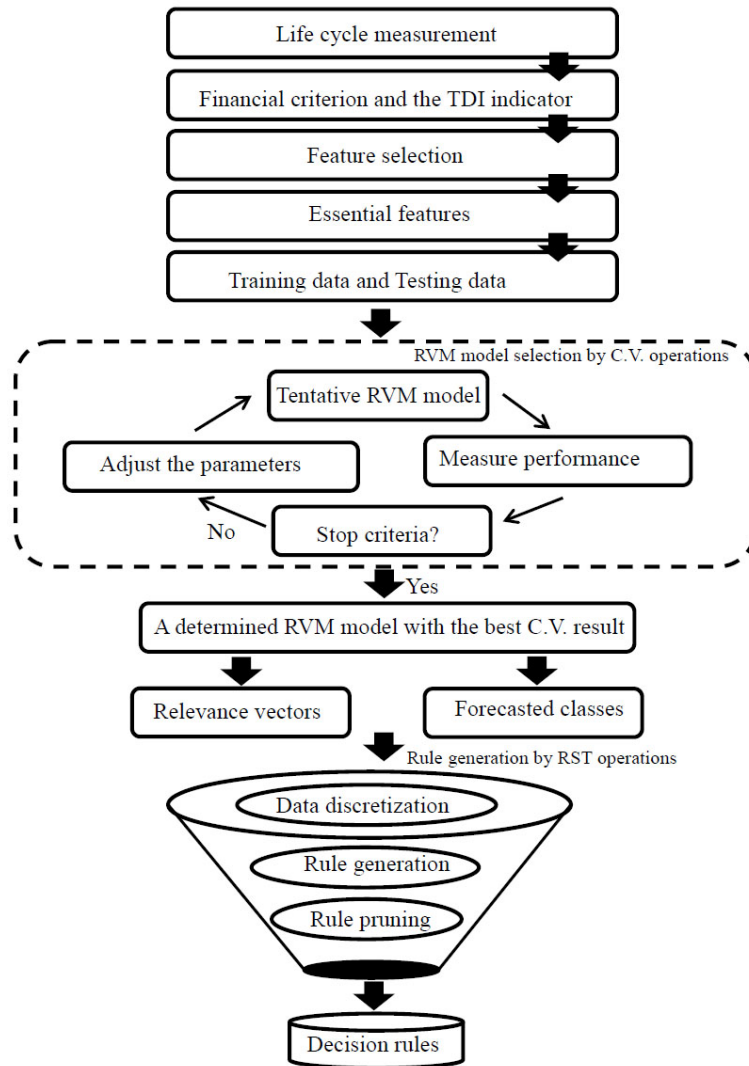**Tab. I** *The criteria in each life cycle stages.*

Fig. 1 *The flowchart of the RVMRS model.*

3.2   Data collection

The investigation collected data of publicly listed electronic industries from 2005 to 2009 as numerical samples. The experimental data included: (1) financial ratios [36] taken from the Taiwan Economic Journal data bank (TEJ) [37] and Taiwan Security Exchange (TSE) [38], (2) the transparency and information disclosure index obtained from Securities and Futures Institute (SFI) [39]. The indicators were divided into five ranks, namely, A^+, A, B, C, and C^-, to illustrate the firms' transparency and disclosure status from superior to inferior. Values from one to

five were utilized to represent the ranking status. Finally, the approach proposed by Anthony and Ramesh [32] was employed to identify the life cycle stages of firms, as shown in Fig. 2.
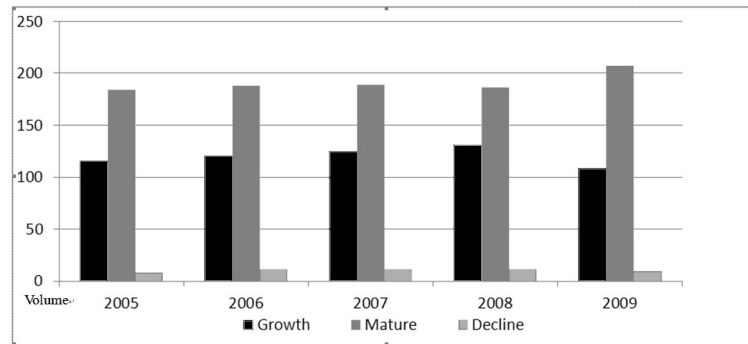
Fig. 2 *The number of corporations of electronic industries in each life cycle stage form 2005 to 2009.*

Although electronic industries in Taiwan have been prosperous and have brought great potential growth in the global capital market, the results of this study show that quite a few corporations have fallen into the decline stage. The investigation used a matched pair research design (1:2); 153 sample corporations were investigated, including the 51 firms that were in the decline stage. Tab. II shows the numbers of firms in terms of their life cycle stages and information disclosure status from 2005 to 2009. It can be observed that corporations in the decline stage have inferior TDI ranking and a weaker governance mechanism. According to Kim et al. [40], corporate governance is essential in explaining the financial performance and market valuation of corporations; even a moderate enhancement of governance structure can considerably increase the performance of firms. In addition, Gompers et al. [41] pointed out that a lower TDI indicator is very much related to lower profit, lower sales growth and higher capital expenditure. Based on the previous work [36], Tab. III presents attributes used in this study to build up the RVMRS model.

Feature selection is a search procedure that identifies a subset of features; it is applied for establishing a robust learning mechanism. Most data used for analysis contains some irrelevant or redundant features that lower the performance generalization of the learning model and bias the investigation findings. In addition, the sequential forward feature selection (SFFS) was used to select the vital features. Starting with an empty set, SFFS repeatedly selects one feature at a time and adds this selected feature into the current feature set. The selected feature has the smallest value of a pre-defined measure among the remaining features. The feature selection procedure is conducted until a maximum of measure has been obtained. The measure used in this study is the recognition rate represented by Eq.(1). The selected features were F2, F4, F9, F12, F13, and F16. Figure 3 illustrates the feature selection procedure using the SFFS technique.

		Rating status from one to five (A^+ to C^-)				
Years	Stages	A^+	A	B	C	C^-
2005	Growth	2	27	58	26	3
	Mature	9	50	99	25	1
	Decline	1	1	2	1	3
2006	Growth	6	33	50	29	3
	Mature	4	56	110	17	1
	Decline	3	2	2	1	3
2007	Growth	2	33	62	26	2
	Mature	8	69	95	16	1
	Decline	2	1	2	1	5
2008	Growth	4	27	62	33	5
	Mature	6	65	96	16	3
	Decline	2	2	1	0	6
2009	Growth	3	25	49	28	4
	Mature	10	68	109	18	2
	Decline	0	1	2	0	7

Tab. II *The number of firms in terms of stages of life cycle and information disclosure status from 2005 to 2009.*

Features	Explanation
F1	Working capital to total assets
F2	Total debt to total assets
F3	Current assets to current liabilities
F4	Operating income to total assets
F5	Net income to total assets
F6	Cash flow to total debt
F7	Cash flow to sales
F8	Sales to total assets
F9	Cash to total assets
F10	Inventory to sales
F11	Current assets to total assets
F12	Long term debt to total assets
F13	Operating income to sales
F14	Current assets to sales
F15	Net income to (total assets – total liabilities)
F16	Transparency and information disclosure index

Tab. III *Definitions of attributes.*

An interesting finding is that the governance indicator (TDI) was selected. The TDI is useful in forecasting corporations that are falling into the decline stage. Prior investigations have pointed out that well-established corporate governance plays a crucial role in alleviating the opportunistic behaviors of management, thus
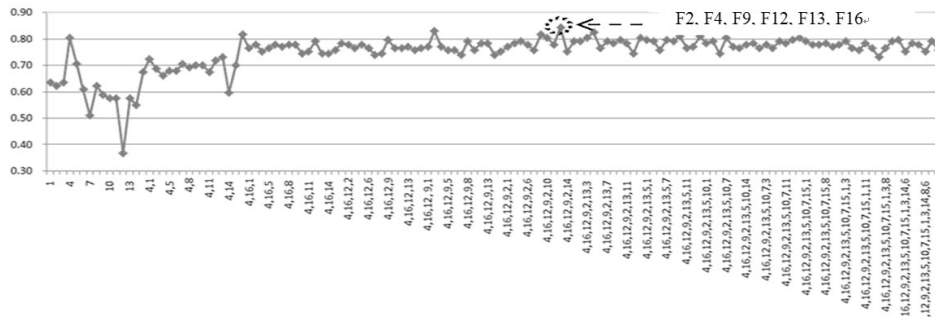
Fig. 3 *Features selected by the SFFS approach..*

enhancing a firm's financial reporting quality and increasing firm value [42, 43]. Transparency and disclosure are most essential in corporate governance. Higher transparency and better disclosure can eliminate information asymmetry as well as mitigate the agency problem [44]. The public sector in Taiwan issued the Corporate Governance Best-Practice Principles in 2002 and established a sequential ranking system to measure the transparency and information disclosure of each corporation listed in the Taiwan Stock Exchange (TSE) [45]. Thus, this study further investigated the feasibility of the TDI indicator in predicting the status of a corporation in the decline stage.

Three criteria, namely classification accuracy, sensitivity and specificity, were used to evaluate prediction performance in this study. Sensitivity and specificity incorporate the misclassification cost in prediction effectiveness, and are expressed as Eq. (19) and Eq. (20), respectively.

$$Sensitivity = True\_Positive/\left(True\_Positive + False\_Negative\right) \tag{19}$$

$$Specificity = True\_Negative/\left(False\_Positive + True\_Negative\right) \tag{20}$$

In addition, the SVM models were utilized to evaluate the same numerical data and the results were compared with those obtained from the RVM models. To demonstrate the importance of the TDI indicator, this investigation further examined the performance of the RVM and SVM models with and without the TDI indicator. Tabs. IV to VII depict the empirical results. According to the experimental results, the TDI indicator is crucial in forecasting the decline stage of firms. This finding is consistent with the report provided by Brown and Caylor [46]. Corporations with better governance practices are more profitable, more valuable, less risky, and yield more dividends.

Explanation has been a vital issue of significant concern in the field of machine learning. For black-box models like RVM, the interpretation of classification decisions is crucial for end users [47, 48]. In this study, relevance vectors and forecasted labels provided by RVM were used by the RS mechanism to generate decision rules. Too many rules are hard for users to learn and they make decision-making difficult. The pruning procedure is therefore necessary to select rules according to their strength values. Two basic indicators, namely testing accuracy and testing

	RVMRS models	SVM models
	Accuracy (%)	Accuracy (%)
CV1	86.7	83.33
CV2	90	83.33
CV3	86.7	86.67
CV4	86.7	86.67
CV5	90	83.33
Average	88	84.67

Tab. IV *The classification accuracy with the TDI indicator.*

	RVMRS models	SVM models
	Accuracy (%)	Accuracy (%)
CV1	80	80
CV2	73.3	76.67
CV3	76.7	80
CV4	83.3	80
CV5	86.7	76.67
Average	80	78.67

Tab. V *The classification accuracy without the TDI indicator.*

	RVMRS models		SVM models	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
CV1	90	85	80	85
CV2	81.8	94.7	81.8	84.21
CV3	88.9	85.7	88.9	85.7
CV4	80	90	80	90
CV5	90	90	80	85
Average	86.1	89.1	82.14	85.98

Tab. VI *The sensitivity and specificity with the TDI indicator.*

coverage, are often employed to indicate the performance of RST. This study used the summation of testing accuracy and testing coverage (STATC) to evaluate the performance of rules derived from RST. Tab. VIII states the outcome of different pruning criteria in the CV-5 dataset. According to Tab. VIII, we can observe that the dataset with the pruning policy "pruning rules with strength values below 2" resulted in the best STATC. The top 5 informative decision rules are listed in Tab. IX. According to the decision rules, the corporations that fell into the decline stage had higher liabilities, lower profitability and a weaker corporate governance structure. Moreover, the decline of a corporation deteriorated its risk absorbing ability and led to critical financial distress. We can expect that the global economic depression or financial tsunami aggravated this decline situation. Layoff is the most

	RVMRS models		SVM models	
	Sensitivity(%)	Specificity (%)	Sensitivity(%)	Specificity(%)
CV1	80	80	70	85
CV2	72.7	73.7	72.7	78.95
CV3	73.3	76.2	77.78	80.95
CV4	80	85	80	80
CV5	80	90	70	75
Average	77.21	81	74.1	79.98

Tab. VII *The sensitivity and specificity without the TDI indicator.*

Pruning criteria	Number of rules	Testing accuracy (%)	Testing coverage (%)	STATC
Initial rule base	57	79	100	1.79
Pruning rule with rule strength below one	43	85	96	1.81
Pruning rule with rule strength below two	35	89	93	1.82*
Pruning rule with rule strength below three	21	92	83	1.75

∗ The pruning criterion with the largest STATC value in Tab. VIII.

Tab. VIII *The testing accuracy, testing coverage and STATC with different pruning policies.*

effective way to change the financial structure of a corporation and reduce its operational costs so that it can survive in a highly competitive NEF environment. This may explain the higher layoff rate in Taiwan during the financial tsunami. Unfortunately, the cost down process typically has side effects: it increases the unemployment rate and causes harsh turbulence to the financial market. Thus, we can hope and expect that the public sectors would take these effects into consideration in formulating future policies that ensure a stable economic environment.

4. Conclusions and future research

This investigation proposed a financial early warning (RVMRS) mechanism incorporating RVM and RS approaches to forecast the decline stage of a corporation and provide rules for decision makers. The problem domain was derived from the life cycle theory and risk management. Numerical data collected from electronic industries were employed to examine the feasibility and performance of the designed RVMRS model. Empirical results revealed that the proposed RVMRS model can not only predict the decline stage of firms with accuracy, but it can also provide rules for decision makers. According the findings of this investigation, corporations in the decline stage usually have higher debt structure, lower prof-

RuleNo.	Strength	Rules
ℜI	57	If the F4 is between 0.079 and 0.125, F9 is between 0.037 and 0.217, and F16 is 1, then the corporate is in the non-decline stage.
ℜII	42	If the F2 is between 0.265 and 0.312, and F16 is 4, then the corporate is in the decline stage.
ℜIII	38	If the F4 is between 0.125 and 0.178, F9 is between 0.037 and 0.217, and F16 is 2, then the corporate is in the non-decline stage.
ℜIV	29	If the F2 is between 0.313 and 0.378, F12 is between 0.147 and 0.301, and F16 is 4, then the corporate is in the decline stage.
ℜV	21	If the F2 is between 0.265 and 0.312, F9 is between 0.002 and 0.079, and F13 is between 0.05 and 0.09, then the corporate is in the decline stage.

Tab. IX *Top five rules provided by the RVMRS model.*

itability, and improper operating efficiency. On the other hand, corporations in the non-decline stage have higher profitability, suitable capital structure, significantly free cash flow, and better operating performance. Thus, corporations in the non-decline stage own stronger risk-buffering capability than corporations in the decline stage. Corporations in the decline stage incline to financial distress; and a warning mechanism is helpful to decision makers in stabilizing capital structures of corporations. In addition, this study examined the usefulness of information and disclosure index (TDI) indicator, which is closely related to corporate governance, in predicting the decline stage of corporations. The empirical results indicated that the corporations in the non-decline stage usually own stronger governance structure and are profitable. This outcome is consistent with the prior study of Brown and Caylor [49]. This represents a significant contribution, since the timely forecast of decline in electronic industries benefits both the development of firms and the capital environment.

This study is the first to predict and analyze the decline stage of electronic firms in Taiwan. Future research might employ data of other industries or from other countries to examine the feasibility of the proposed RVMRS model. Another possible research direction would be to apply other classifiers and rule generation mechanisms to improve the analysis and forecasting performance.

Acknowledgement

References

[1] Granlund M., Taipaleenmäki J.: Management control and controllership in new economy firms-a life cycle perspective. Manage. Account. Res. **16**, 2005, pp. 21-57.

[2] Clark P. J., Neill S.: Net value. In: Valuing Dot-Com Companies—Uncovering the Reality Behind the Hype. Amacom, New York, 2001.

[3] Greiner L.: Evolution and revolution as organizations grow. Harvard Bus. Rev. **50**, 1972, pp. 37–46.

[4] Adizes I.: Organisational passages: Diagnosing and treating life-cycle problems in organizations. Organ. Dyn. **8**, 1979, pp. 3–25.

[5] Owen S., Yawson A.: Corporate life cycle and M&A activity. J. Bank. Financ. **34**, 2010, pp. 427-440.

[6] Miller D., Friesen P.: A longitudinal study of the corporate life cycle. Manage. Sci. **30**, 1984, pp. 1161-1183.

[7] Black E. L.: Lifecycle impacts on the increment value-relevance of earnings and cash flow measure. J. Financ. Statement Anal. **4** , 1998, pp. 40-56.

[8] Beaver W. H.: Financial ratios as predictors of failure, empirical research in accounting: Selected studies. J. Account. Res. **4**, 1966, pp. 179–199.

[9] Altman E. I.: Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. J. Financ. **23**, 1968, pp. 589-609.

[10] Altman E. I., Haldeman R. G., Narayanan P.: Zeta analysis. J. Bank. Financ. 1977, pp. 29–51.

[11] Zavgern C. V.: The Prediction of Corporate Failure: The State of the Art. J. Account. Literature, **2**, 1983, pp. 1-38.

[12] Kantardzic, M. 2002. Data Mining: Concepts, Models, Methods, and Algorithms. IEEE Press, Piscataway.

[13] Samsudin R., Saad P., Shabri A.: A hybrid GMDH and least squares support vector machines in time series forecasting. Neural Netw. World, **3**, 2011, pp. 251-268.

[14] Darwish A., Abraham A.: The use of computational intelligence in digital watermarking: Review, challenges, and new trends, Neural Netw. World. **4**, 2011, pp. 277-297.

[15] Sulaiman S., Shamsuddin S. M., Abraham A., Sulaiman S.: Intelligent web caching using machine learning methods. Neural Netw. World, **5**, 2011, pp. 429-452.

[16] Zhang H., Wu W.: Convergence of split-complex backpropagation algorithm with momentum Neural Netw. World, **1**, 2011, 75-90.

[17] Kulkarni S.: Fingerprint feature extraction and classification by learning the characteristics of fingerprint patterns, Neural Netw. World. **3**, 2011, pp. 219-226.

[18] Tipping M. E.: Sparse Bayesian learning and the relevance vector machine. J. Mach. Learn. Res. **1**, 2001, pp. 211-244.

[19] Wei L.Y., Yang Y.Y., Nishikawa R.M., Wernick M.N., Edwards A.: Relevance vector machine for automatic detection of clustered microcalci?cations. IEEE Trans. Med. Imaging **24**, 2005, pp.1278–1285.

[20] Williams O., Blake A., Cipolla R.: Sparse Bayesian learning for ef?cient visual tracking, IEEE Trans. Pattern Anal. Mach. Intell. **27**, 2005, pp.1292–1304.

[21] Demir B., Erturk S.: Hyperspectral image classi?cation using relevance vector machines. IEEE Geosci. Remote Sens. Lett. **4**, 2007, pp.586–590.

[22] Pawlak Z., Skowron A.: Rudiments of rough sets. Inform. Sciences **177**, 2007, pp.3–27.

[23] Pudil P., Novovicova J., Kittler J.: Floating search methods in feature selection. Pattern Recogn. **15**, 1994, pp. 1119–1125.

[24] Majumder S. K., Ghosh N., Gupta P. K.: Relevance vector machine for optical diagnosis of cancer. Lasers Sur. Med. **36**, 2005, pp. 323–333.

[25] Han M., Zhao Y.: Dynamic control model of BOF steelmaking process based on ANFIS and robust relevance vector machine. Expert Syst. Appl. **38**, 2011, pp. 14786-14798.

[26] Pawlak Z.: Rough sets. Int. J. Comp. Inform. Sci. **11**, 1982, pp. 341–356.

[27] Pawlak Z.: Roughsets, Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Dordrecht, 1991.

[28] Mark B., Munakata T.: Computing artificial intelligence and information technology. Eur. J. Oper. Res. **136**, 2002, pp. 212–229.

[29] Pawlak Z.: Rough sets and intelligent data analysis. Inform. Sciences, **147**, 2002, pp. 1–12.

[30] Wang X., Yang J., Jensen R., Liu X.: Rough set feature selection and rule induction for prediction of malignancy degree in brain glioma. Comput. Meth. Prog. Bio. **83**, 2006, pp. 147–156.

[31] Dimitras A. I., Slowinski R., Susmaga R., Zopounidis C.: Business failure prediction using rough sets. Eur. J. Oper. Res. **114**, 1999, pp. 263–280.

[32] Anthony J. H., Ramesh K.: Association between accounting performance measures and stock prices. J. Account. Econ. **15**, 1992, pp. 203–227.

[33] Sotoca J. M., Pla F.: Supervised feature selection by clustering using conditional mutual information-based distances. Pattern Recogn. **43**, 2010, pp. 2068-2081

[34] Salzberg S.: On Comparing Classifiers: Pitfalls to Avoid and a Recommended Approach. Data Min. Knowl. Disc. **1**, 1997, pp. 317-328.

[35] Davis C. P., Cleven D., Balish E., Yale C. E.: Bacterial association in the gastrointestinal tract of beagle dogs. Appl. Environ. Microb. **34**, 1977, pp. 194-206.

[36] Bose, I.: Deciding the financial health of dot-coms using rough sets. Inform. Manage. **43**, 2006, pp. 835-846.

[37] Taiwan Economic Journal Co., Ltd. < http://www.tej.com.tw/twsite/>

[38] Taiwan Stock Exchange Co.< http://www.twse.com.tw/ch/index.php>

[39] Taiwan Securities and Futures institute Co. < http://www.sfi.org.tw/newsfi/chinese.asp>

[40] Kim W., Black B. S., Jang H.: Does corporate governance predict firms' market values? Evidence from Korea. J. Law, Econ. Organ. **22**, 2006, pp. 366–413.

[41] Gompers P. A., Ishii J. L., Metrick A.: Corporate governance and equity prices. Q. J. Econ. **118**, 2003, pp. 107–155.

[42] Denis D. K., McConnell J.: International corporate governance. J. Financ. Quantit. Anal. **38**, 2003, pp. 1–36.

[43] Bhagat D., Bolton B.: Corporate governance and firm performance. J. Corp. Financ. **14**, 2008, pp. 257–273.

[44] Patel S. A., Balic A., Bwakira L.: Measuring transparency and disclosure at firm-level in emerging markets. Emerging Markets Rev. **3**, 2002, pp. 325-337.

[45] Chi L. C.: Do transparency and disclosure predict firm performance? Evidence from the Taiwan market. Expert Syst. Appl. **36** (8), 2009, pp. 11198-11203.

[46] Brown L. D., Caylor M. L.: Corporate governance and firm performance. J. Account. Public Pol. **25**, 2006, pp. 409–434.

[47] Wyatt J.: Nervous about artificial neural networks? Lancet, **346**, 1995, pp. 1175–1177.

[48] Fung G., Sandilya S., Rao R.: Rule extraction from linear support vector machines, in: Proceedings of the Eleventh SIGKDD International Conference on Knowledge Discovery and Data Mining, 2005.

[49] Brown, L. D., Caylor, M. L.: Corporate governance and ?rm performance. J. Account. Public Pol. **25**, 2006, pp.409-434.