International Journal of Industrial Engineering and Operations Management (IJIEOM)
Volume. 1, No. 1, pp. 78-96, May 2019

IEOM Society International

International Journal of Industrial Engineering and Operations Management (IJIEOM)

Volume 1, No. 1, May 2019
pp. 78 - 96

Bankruptcy Prediction for Japanese Corporations using Support Vector Machine, Artificial Neural Network, and Multivariate Discriminant Analysis

Masanobu Matsumaru
Research Institute for Engineering, Kanagawa University, Kanagawa, Japan
pt120916cz@kanagawa-u.ac.jp

Takaaki Kawanaka
Institute for Innovation in International Engineering Education
Graduate School of Engineering, The University of Tokyo, Tokyo, Japan
kawanaka@cce.t.u-tokyo.ac.jp

Shoichi Kaneko
Faculty of Business Administration, Yamanashi Gakuin University, Yamanashi, Japan
shoichi@ygu.ac.jp

Hideki Katagiri
Department of Industrial Engineering and Management,
Faculty of Engineering, Kanagawa University, Kanagawa, Japan
katagiri@kanagawa-u.ac.jp

ABSTRACT

This study predicted the bankruptcy risk of companies listed in Japanese stock markets for the entire industry and individual industries using multiple discriminant analysis (MDA), artificial neural network (ANN), and support vector machine (SVM) and compared the methods to determine the best one. The financial statements of the companies listed in the Tokyo Stock Exchange in Japan were used as data. The data of 244 companies that went bankrupt between 1991 and 2015 were used. Additionally, the data of 64,708 companies that did not go bankrupt between 1991 and 2015 (24 years) were used. The data was acquired from the Nikkei NEEDS database. It was found from the results of empirical analysis that the SVM is more accurate than the other models in predicting the bankruptcy risk of companies. In the ANN analysis and MDA, bankruptcy prediction could be made accurately only for some individual industries. In contrast, the SVM could predict the bankruptcy risk of companies almost perfectly for either entire and individual industries. This bankruptcy prediction model can help customers, investors, and financiers prevent losses by focusing on the financial indicators before finalizing transactions.

ARTICLE INFO

Received
Feb 18, 2019
Received in revised form
May 03, 2019
Accepted
May 10, 2019

KEYWORDS
Bankruptcy, artificial neural networks (ANN), support vector machines (SVM), multiple discriminant analysis (MDA), empirical analysis

1. Introduction

Bankruptcy is intricately related to the economic environment. When the economy is good, the number of bankrupt companies decreases, and when the economy is bad, the number of bankrupt companies increases. Thus, economic dynamics impact the rate of corporate bankruptcy. Naturally, this has been the case with the Japanese economy. When we look back on the history of the Japanese economy in the past 30 years, we observe that it recorded the highest rate (38,915 yen) in the average history of the Tokyo Stock Exchange (TSE) in 1989, and the economy had been steady until that time. The business environment was at its best until this time; therefore, there were very few bankrupt companies. However, the Berlin Wall collapsed at the end of 1989, the Cold War ended, and a new era of world competition began. A new business environment emerged, and it became difficult to manage businesses. Between 1990 and 2000, during the 10 year Heisei recession, corporate bankruptcies did not increase in Japan; however, following the collapse of the bubble economy in 2001, corporate bankruptcies increased due to bad debts of banks and the deflationary recession. After that, the Japanese economy recovered moderately, and the number of bankruptcies decreased. The Japanese economy gradually recovered and the number of bankruptcies decreased until the major recession in 2007 due to the subprime loan crisis, when the number began increasing again. After this subprime loan crisis, the Japanese economy stagnated for a while. Although the US default crisis occurred in October 2013, Japanese companies were not significantly affected, and bankruptcies did not increase. In recent years, the Japanese economy has been gradually recovering due to economic developments incorporating new services such as Fintech. Japanese companies have been improving their corporate performance and financial structure because of the depreciation of the yen and the gradual recovery of the overall economy. Compared with the internal reserves of companies at the end of FY2011 and at the end of FY2017 just before the second Abe administration was established, the internal reserves of companies were said to have increased by about 164 trillion yen. The internal reserve amount of the company has reached the maximum. On the other hand, the Japanese economy today is said to be a bubble whose collapse can occur at any time. The possibility that corporate bankruptcy will occur due to the collapse of the bubble is increasing. Although the number of bankruptcies has decreased, the likelihood of a company going bankrupt remains a highly important management subject.

This study predicted the bankruptcy risk of companies listed in the Japanese stock markets for the entire industry and individual industries using the Multiple Discriminant Analysis (MDA), artificial neural networks (ANN), and support vector machines (SVM), and compared the methods to determine which method most accurately predicts bankruptcy risk. Bankruptcy prediction has drawn a lot of research interests in previous literature. A bankruptcy discrimination research was first conducted by Beaver (1966). Altman (1968) classified companies into groups using the MDA to distinguish between bankrupt and non-bankrupt companies, and further developed Beaver's research. Deakin (1972) pointed out that most of such assumptions were not valid because the financial data are independent variables. Therefore, it can be inferred that statistical methods have limitations in achieving effectiveness and validity. Holland (1975) presented genetic algorithm (GA) in adaptation in natural and artificial systems. Ohlson (1980) employed first logistic regression. Odom and Sharda (1990) first applied ANN to bankruptcy prediction. Altman, Marco, and Varetto (1994) compared the accuracy of linear discriminant formula (LDA) and logit analysis (LA). Boritz and Kennedy (1995) compared ANN against traditional bankruptcy prediction techniques such as discriminant analysis, logit, and probit. Wilson and Sharda (1994)) demonstrated that the ANN performs significantly better than discriminant analysis at predicting firm bankruptcies. Zurada et al. (1994) showed that multilayer feedforward networks are often used for modeling complex relationships between the data sets. Back et al. (1996) focused on the MDA, LA, and GA. Jo and Han (1996) used the MDA, ANN, and case-based forecasting system. Jo, et al. (1997) used MDA and case-based forecasting to predict corporate bankruptcy. Barniv et al. (1997) used the ANN, multi-state ordered logit, and nonparametric multiple discriminant analysis (NPDA). Bell (1997) modeled regulator decision using LA and ANN. Zhang et al. (1999) presented a general framework for understanding the role of the ANN, and compared it with the LA. Cheng et al. (2006) combined the approaches of ANN learning and LA. Chung et al. (2008) utilized the MDA and ANN to create an insolvency predictive model in New Zealand. Tseng and Hu (2010) used LA, quadratic interval LA, backpropagation multi-layer perceptron (MLP), and radial basis function network (RBFN). Mohatab et al. (2011) used the ANN, GA and MDA in Iranian evidence. Jeong et al. (2012) suggested incorporating the generalized additive model (GAM) and GA into the ANN. Lee and Choi (2013) presented a multi-industry investigation of the bankruptcy of Korean companies using back-propagation neural network (BNN) and compared BNN and the MDA. Kasgari et al. (2013) utilized the multilayer perceptron (MLP) of ANN, comparing it with a probit analysis. Iturriaga and Sanz (2015) developed a model of ANN that combines multilayer perceptrons and self-organizing maps. Inthachot, et al. (2016) investigated the use of ANN and GA for predicting Thailand's SET50 index trend. Cortes and Vapnik (1995) presented the support-vector network as a new learning machine for two-group classification problems. Vapnik (1998) introduced SVM in the book of statistical learning theory. This book has three parts, and support vector estimation is one part of this book. Weston et al. (1999) described how the SVM technique of solving linear operator equations can be applied to the problem of density estimation. Crammer and Singer (2001) used SVM on the algorithmic implementation of multiclass kernel. Chang and Lin (2001) have been actively developing an SVM package since the year 2000 to help users to easily apply SVM to their applications. Hsu and Lin (2002) presented

the decomposition method for solving SVM. Kim (2003) studied the feasibility of using SVM to predict the stock price index. Karatzoglou et al. (2004) presented an S4 Package for Kernel method in R. Min and Lee (2005) used SVM, and compared its performance with those of the MDA, LA, and three-layer fully connected back-propagation neural networks (BPN). Shin et al. (2005) showed that the proposed classifier of the SVM approach outperforms BPN. Min et al. (2006) have dealt with the integration of GA and SVM. Boyacioglu et al. (2009) applied ANN, SVM and multivariate statistical methods to bank failure in Turkish. Yang and Ji (2011) combined the partial least squares-based (PLS-based) feature selection with SVM for information fusion. Kim (2011) used MDA, LA, ANN, and SVM for hotel bankruptcy prediction. Chen, H.Y. (2011) applied particle swarm optimization (PSO) to obtain suitable parameter settings for SVM. Chen, M.Y. (2011) utilized the traditional statistical, Decision Tree and ANN in Taiwan Stock Exchange Corporation (TSEC) listed companies. Fatih (2013) utilized ANN and SVM in Turkish bank data. DONG et al. (2015) investigated the impact of data comparability on the performance of SVM models for credit scoring. Matsumaru et al. (2018) used the decision tree method and SVM for bankruptcy prediction on the entire industry and 11 industries in Japan stock market.

In this research, the computer used to run all simulations was a PC (OS: Windows 10 Home Premium 64 bit) equipped with Intel (R) Core (TM) i5-8400 CPU@2.8 GHz and 8.0 GB memory. The software used to perform the simulation, predict, and conduct the overall study is the open source statistical software R version 3.5.1.

2. Hypothesis Setting and Evaluation Criteria

We propose three hypotheses as basic ideas to advance research, and take research methods to verify them. The first hypothesis is that forecasts can predict more accurately using many financial ratios than with fewer ones. The second hypothesis is that forecasts of one industry type as the entire industry without classification can be predicted more accurately than if classified into individual industries. The third hypothesis is that prediction using SVM can predict more accurately than prediction using other methods. The explanatory variable of the first hypothesis uses 23 financial ratios and 115 financial ratios. The 23 financial ratios are used the results obtained from previous studies by Koshika and Matsumaru et al. (2014). The 115 financial ratios are the number of results from collecting as many financial ratios as possible from the Nikkei NEEDS database. The financial statements of the listed companies in TSE are used as data. The data of 244 companies that went bankrupt between 1992 and 2015 are used. Additionally, the data of 64,708 companies that did not go bankrupt between 1992 and 2015 are used. Both the 23 financial ratios and 115 financial ratios are listed in the Appendix.

To verify the second hypothesis, two types of bankruptcy prediction are carried out: entire industry and individual industries. Individual industries were classified into 33 industries. The industry classification at this time is based on TOPIX Sector financial ratios. TOPIX Sector financial ratios list 33 industries that the TSE classified based on the business category classification set by the "Securities Code Council". Of the 33 industries, banks are excluded from this research because they are considered special compared to other industries. Of the 32 remaining industries, bankruptcy occurred in 22 industries, but bankruptcy did not occur in the remaining 10 industries. These industries are Fishery, Agriculture and Forestry, Pharmaceutical, Oil and Coal Products, Nonferrous Metals, Precision Instruments, Electric Power and Gas, Land Transportation, Marine Transportation, Securities, and Commodities Futures, Insurance. Of the 22 industries, 10 industries with many bankrupt companies are handled as independent industries; specifically, Construction, Textiles and Apparels, Machinery, Electric Appliances, Information and Communication, Wholesale Trade, Retail Trade, Other Financing Business, Real Estate and Services. In the remaining 12 industries, there are few bankrupt companies, therefore, 31 bankrupt companies belonging to 12 industries are treated as one industry, named “other industry”. As a result, the number of industries contemplated in this study is 11. The relationship between the 11 industries and the 244 bankrupt companies is as follows: The data of 244 bankrupt companies are classified under 11 industries: construction industry (43 companies), real-estate industry (34 companies), services industry (20 companies), retail industry (19 companies), electrical equipment industry (17 companies), machinery industry (16 companies), wholesale industry (15 companies), other financial services industry (13 companies), textiles and apparels industry (9 companies), information and telecommunications industry (8 companies), and other industries (31 companies). To verify the third hypothesis, three methods of MDA, ANN, and SVM are used. These three methods are described in the next chapter.

The fitness function f used as evaluation criteria herein is expressed in Eq. (2.1). The fitness function is defined as the number of successes in the correct classification of bankrupt and non-bankrupt companies. This function utilizes the financial ratios of both correct and incorrect predictions:

$$f = \frac{T_p + F_n}{T_p + T_n + F_p + F_n} \tag{2.1}$$

In the above, true positive (T_p) is the number of bankrupt companies that the rule rightly classifies as bankrupt, false positive (F_p) is the number of non-bankrupt companies that the rule erroneously classifies as bankrupt, true negative (T_n) is the number of non-bankrupt companies that the rule rightly classifies as non-bankrupt, and false negative (F_n) is the number of bankrupt companies that the rule erroneously classifies as non-bankrupt. Furthermore, the evaluation is also performed using the evaluation index such as the precision and the recall. In the case of two classifications of bankrupt companies and non-bankrupt companies, it is a 2 × 2 confusion matrix. The elements of this confusion matrix are T_p, F_p, T_n and F_n. The precision is defined as follows:

$$Precision = \frac{T_p}{T_p + F_p} \tag{2.2}$$

The recall is defined as follows.

$$Recall = \frac{T_p}{T_p + F_n} \tag{2.3}$$

The model performance is judged to be higher as the recall and the precision are higher.

3. Theoretical Framework

A brief description of each tool that is used in this study is given subsequently.

3.1 Multiple discriminant analysis (MDA)

The MDA can be used for selecting bankrupt companies using statistical theory to build a bankruptcy prediction model. The MDA is a model developed by Altman (1968) and is a method for making an objective judgment based on observed values of n variables. Classification is generally performed by a discriminant function in the form of a linear combination of independent variables. That is, in the linear discriminant function, such discrimination is performed on the basis of the linear form of n variables $x_1, x_{2,\cdots}, x_n$. In the case of bankruptcy prediction analysis, the classification is binary and the linear discriminant function is as follows:

$$z = a_1x_1 + a_2x_2 + \cdots\cdots + a_ix_i \tag{3.1}$$

where, z represent a discriminant score, a_i represent discriminant coefficients, and x_i represent financial ratios ($i = 1,2,\cdots,n$).

The discriminant coefficients are derived in such a way as to minimize the possibility of misclassification. The discriminant coefficients are derived by minimizing the ratio of between group and within-group variances. Use the discriminant scores to classify objects into the appropriate bankruptcy group (Group 1) and non-bankruptcy group (Group 0).

3.2 Artificial neural networks (ANN)

3.2.1 Outline of neural network

The neural network is a problem optimization method modeling the neural circuit of the human brain. The neural network is trained with or without a teacher signal. A teacher signal is used when the solution is already known, as is the case here. No teacher signal is used to classify given arbitrary data. Additionally, it is classified as a hierarchical type and/or mutually coupled type network by the neuron connection method. With supervised signals, the hierarchical network approach is a perceptron, using the back-propagation algorithm. Basically, the perceptron is trained by providing the correct answer to the output layer as a teacher signal, and the error between the teacher signal and the output signal is calculated and back-propagated to modify the weights of the connections among neurons in a repetitive fashion until the algorithm is stopped.

3.2.2 Bankruptcy discrimination model using ANN

The bankruptcy discrimination model using ANN predicts bankruptcy if the value of the output layer is "1". In this research, the number of learning steps is 100,000, and a bankruptcy discrimination model is constructed using three layers: the input layer, the intermediate (hidden) layer, and the output layer; the model is trained using the backpropagation method. Neurons in the input layer receive/represent the financial ratios, and the number of neurons in the middle layer changes for each index number. The number of neurons in the output layer is 1. The output unit displays a true/false relationship (binary 1 for bankrupt, and 0 for non-bankrupt). Additionally, the ANN model uses variables. The number of neurons in the input layer is determined from among one to 23 of the number of financial ratios, based on the in-sample data and out of sample results. The indicator with the highest discrimination power and reliability is

determined; the ANN model is determined as well. A sigmoid function was used as the activation function for the computing neurons, and the open source statistical software R was used in numerical experiments.

3.3 Support vector machines (SVM)

An SVM, presented as machine learning by Vapnik (1998), is a data analysis method that mainly deals with classification and regression problems. An SVM is a high dimensional hypothesis space that can be linearly separated, and it can be understood as follows as it is a method that follows a linear approach. We define the training data sets as $(x_{1,}y_1), (x_{2,}y_2), \cdots\cdots, (x_{n,}y_n)$, where $\boldsymbol{x} = (x_1, x_2, \cdots\cdots, x_p)^T$ is the feature vector of the individual and y is the objective variable that is a numerical value in the regression problem and it is the label of the class in the classification problem. The data consists of a pair of feature vectors $x_i \in \boldsymbol{R}^n, i$=1, 2, $\cdots l$ and class $y_i \in \{-1,1\}$. $\boldsymbol{x}_i$ is the financial indicator of company i, such that y_i=-1 represents bankruptcy and y_i=1 indicates non-bankruptcy. For linear regression and linear discrimination problems, the following linear model is used:

$$\mathrm{y} = \boldsymbol{w}^T\boldsymbol{x} + b \tag{3.2}$$

Positive and negative samples are separated by the hyperplane H_0: $\boldsymbol{w}^T\boldsymbol{x} + b = 0$.

SVM determines the coefficient that maximizes the margin, and performs discrimination as follows:

$$\mathrm{y} = \begin{cases} 1, & \text{if} \quad \boldsymbol{w}^T\boldsymbol{x} + b \geq 1 \\ -1, & \text{if} \quad \boldsymbol{w}^T\boldsymbol{x} + b \leq -1 \end{cases} \tag{3.3}$$

The interval between the straight lines $\boldsymbol{w}^T\boldsymbol{x} + b = 1$ and $\boldsymbol{w}^T\boldsymbol{x} + b = -1$ is the margin M; maximizing the margin M is the problem of maximizing: $M = \min\left(\frac{|y_i(\boldsymbol{w}^T x_i + b)|}{\|w\|}\right) = \frac{2}{\|W\|}$.

The maximization of the margin M is equal to the minimization of $\|w\|$; therefore, the SVM problem for obtaining hyperplane H_0 that maximizes the margin is formulated as the following quadratic programming problem:

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} \\ &\text{Subject to} \quad y_i(\boldsymbol{w}^T x_i + b) \geq 1; i = 1,2, \cdots, n \end{aligned} \tag{3.4}$$

If it is possible to convert the specimen data that cannot be linearly separated into another space that enables linear separation, then we can use an SVM so that the boundary between the groups becomes a hyperplane in the space of the conversion destination. Therefore, the space in which the original sample data exists is called the input space, and the space of the destination is called the feature space($\emptyset(x_i)$). After input sample x_i is transformed into the sample in the feature space, the SVM can be applied to the sample because the sample is linearly separated in the feature space. Corresponding to Eq. (3.4), the discriminant hyperplane in feature space $\emptyset(x_i)$ becomes the optimal solution for the following minimization problem:

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} \\ &\text{Subject to} \quad y_i(\boldsymbol{w}^T\emptyset(x_i) + b) \geq 1; i = 1,2, \cdots, n \end{aligned} \tag{3.5}$$

Relaxation variable $\varepsilon_i \geq 0 (i = 1,2, \cdots, n)$ is introduced so that sample data that does not satisfy the constraint condition in Eq. (3.5) may exist. The constraint condition of Eq. (3.5) is relaxed as follows:

$$y_i(\boldsymbol{w}^T\emptyset(x_i) + b) \geq 1 - \varepsilon_i \tag{3.6}$$

Under the relaxed constraints of Eq. (3.6), soft margin SVM is formulated as follows:

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C\sum_{i=1}^{n}\varepsilon_i \\ &\text{Subject to } y_i(\boldsymbol{w}^T\emptyset(x_i) + b) \geq 1 - \varepsilon_i \\ &\varepsilon_i \geq 0; i = 1,2, \cdots\cdots, n \end{aligned} \tag{3.7}$$

Where, $\sum_{i=1}^{n}\varepsilon_i$ is the upper limit value of sample data misclassified, and C (C $\geq$ 0) is a penalty. SVM has several improved forms, one of which is support measure machine (SMM) by the kernel method. SVM by the kernel method is a nonlinear classifier that is represented by a linear function using a kernel function. The purpose of the package kernlab is to provide the R user with a basic kernel functionality in the kernel-based methods. Package kernel is available from CRAN (https://cran.r-project.org/) under the GPL license.

Empirical analysis result

4.1 Results and discussion

4.1.1 Results of entire industries prediction

The results of the entire industries prediction are shown in Table 1. In the table, 0 represents bankruptcy and 1 represents non-bankruptcy. Table 1 shows the table consisting of the fitness function *f*, and the confusion matrix which

The results of the entire industries prediction are shown in Table 1. In the table, 0 represents bankruptcy and 1 represents non-bankruptcy. Table 1 shows the table consisting of the fitness function *f*, and the confusion matrix which have the elements of the numerical values of T_p , F_p , T_n and F_n. In addition, the numerical values of T_p , F_p , T_n and F_n represent the number of companies.

Table 1. Entire industries

	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
MDA	Status	0	43	181	224	99.38	54	170	224	99.27
		1	219	64489	64708		303	64405	64708	
	%	0	19.20	80.80	100.00		24.11	75.89	100.00	
		1	0.34	99.66	100.00		0.47	99.53	100.00	
	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
ANN	Status	0	0	224	224	99.66	73	151	224	99.65
		1	0	64708	64708		78	64630	64708	
	%	0	0.00	100.00	100.00		32.59	67.41	100.00	
		1	0.00	100.00	100.00		0.12	99.88	100.00	
	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
SVM	Status	0	219	5	224	99.99	167	57	224	99.91
		1	0	64708	64708		0	64708	64708	
	%	0	99.77	2.23	100.00		74.55	25.45	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	

It was found from the result that SVM predicts the bankrupt companies more accurately than MDA and ANN. The values of the fitness function *f*, the precision and the recall, using MDA and ANN analysis are less than the values using SVM. From these results, the entire industry will be classified into individual industries subsequently, and research will be further advanced.

4.1.2 Results of each industry prediction

The results of each industry prediction are results are shown in Table 2–Table 12. The results of SVM using 23 financial ratios have already been predicted by Koshika and Matsumaru et al. (2018); however, because for the purpose of comparison with the results gotten from using ANN and MDA, they are included in the table.

4.2 Discussion

We conclude previously from the results of empirical analysis that the SVM is more accurate than the other models in predicting the bankruptcy risk of companies. Here, we discuss the results of empirical analysis using MDA, ANN and SVM of other countries in the world. It clearly states that Odom, M., and Sharda (1990) compared the prediction ability of both neural networks and discriminant analysis in data obtained from Moody's Industrial Manuals, and ANN is more accurate than MDA. Altman（1994）states that the networks have shown significant capacities for recognizing the health of companies, with results that are, in many cases, near or superior to the results obtained through discriminant analysis in the Italian data. Wilson and Sharda (1994) state that neural networks clearly outperformed discriminant analysis in prediction accuracy of both bankrupt and non-bankrupt firms under varying training and testing conditions in the sample obtained from Moody′s Industrial Manuals. Boyacioglu et al. (2009）applied various neural network

techniques, support vector machines and multivariate statistical methods to the bank failure prediction problem in a Turkish case. In the category of neural networks, four different architectures namely multi-layer perceptron, competitive learning, self-organizing map and learning vector quantization are employed. The multivariate statistical methods; namely, multivariate discriminant analysis, k-means cluster analysis and logistic regression analysis are tested. Results show that multi-layer perceptron and learning vector quantization can predict accurately bank financial failures than the multivariate statistical methods. Mohatab et al. (2011) utilized three methods, ANN, GA, and MDA to predict financial health of companies quoted in Tehran Stock Exchange. The result showed that ANN model outperforms the other models captured by two other methods. Kim (2011) reported that the empirical results showed that ANN and SVM were very applicable models in bankruptcy prediction with data from Korean hotels and showed that ANN and SVM with their flexible modelling capability did provide more accurate estimates, leading to higher classification rates than other multivariate statistical methods. Chen, M.Y. (2011) suggests that the SVM could be a more suitable method than the traditional statistical, Decision Tree and ANN techniques, using data collected from 200 Taiwan Stock Exchange Corporation (TSEC) listed companies. Fatih (2013) concludes that both ANN and SVM were promising prediction models as they have achieved over 85% validation prediction performance in the Turkish case. As we have seen above, in the results of empirical analysis using MDA, ANN and SVM around the world, SVM is more likely precisely than other models in predicting bankruptcy risk in a company, similar to the result of this paper. It can be said that it is an accurate bankruptcy forecasting method.

This research predicted the bankruptcy risk of all the industries and individual industries using MDA, ANN, and SVM. We conclude from the results of empirical analysis that the SVM is more accurate than the other models in predicting the bankruptcy risk of companies. We discuss here that the results of empirical analysis are dependent on theoretical advantages of SVM than the methodology of MDA and ANN, not data samples. We would like to say that SVM is a better way to predict bankruptcy than MDA and ANN. In other words, we discuss the advantages of SVM and the disadvantages of MDA and ANN. Regarding the disadvantage of MDA, MDA has the disadvantage that it can't discriminate bankrupt companies and non-bankrupt companies close to boundaries classified by linear discriminant functions because it is linear. SVM overcomes the disadvantages of MDA. SVM does not use data of companies that are far from the boundary and that are clearly identified as bankruptcy or non-bankruptcy, that is, SVM use only the support vector and performs classification. On the other hand, The MDA uses data of all companies including the data that are far from the boundary. The MDA is a theory that calculates the ratio that minimizes the intra-group variation (variance) with the inter-group variation (variance) at a maximum, and calculates the discrimination coefficient so that the ratio is maximized. The MDA is influenced by the value of company data that is far from the boundary. Therefore, The MDA has the disadvantage that misclassification becomes large. On the other hand, the SVM can clearly distinguish between bankrupt and non-bankrupt companies, even for companies close to the border. The reason is that classification is performed using only data near the boundaries, that is, support vectors. In other words, SVM does not use data of companies that are clearly identified as bankruptcy or non-bankruptcy, and performs classification. The SVM calculates the discrimination coefficient by setting the distance between the boundary and the data, that is, the condition that the boundary line and the data should be as far apart as possible, in other words, to maximize the margin M. Therefore, the SVM can prevent erroneous determination. The maximization of the margin M is equal to the minimization of $\|w\|$. Therefore, soft margin SVM is formulated into Eq. (3.7) under the relaxed constraints of Eq. (3.6). The soft margin SVM is formulated in consideration of the condition that the boundary line and the data should be as far apart as possible, and the condition that the number of erroneous determinations should be as small as possible. To be specific, it is formulated such that the sum of the first term, $\frac{1}{2}w^T w$, and the second term, $\sum_{i=1}^{n} \varepsilon_i$, are minimized. At this time, C (C ≥ 0) is introduced as a penalty. Here, the parameter "C" is a parameter representing "to what extent erroneous determination is allowed". If the parameter C is large, it means that misclassification is never permitted. In the case of this study, as described later, C = 10 showed good prediction results. Most of SVM related packages use this soft margin SVM as a standard SVM approach.

Table 1 showed the table consisting of the fitness function *f*, and the confusion matrix which have the elements of the numerical values of T_p, F_p, T_n and F_n for entire industry prediction. At first, we evaluate the model. Basically, it is important that the model predict precisely the bankruptcy company as the bankruptcy and no-bankruptcy company as the non-bankruptcy precisely. From such a point of view, we focus on T_p and T_n among T_p, T_n, F_p and F_n. T_p is the number of bankrupt companies that the rule precisely predicts as bankrupt, and F_n is the number of non-bankrupt companies that the rule precisely predicts as non-bankrupt. Using the MDA, T_p is 19.20 % and T_n is 99.66 % for 23 financial ratios, while T_p is 24.11 % and T_n is 99.53 % for 115 financial ratios.

Using the ANN, T_p is 0.00% and T_n is 100.00% for 23 financial ratios, while T_p is 32.59% and T_n is 99.88% for 115 financial ratios. Using the SVM, T_p is 97.77 % and T_n is 100.00 % for 23 financial ratios, while T_p is 74.55 % and T_n is 100.00% for 115 financial ratios. These results indicate that the SVM can predict bankrupt companies precisely and is

more accurate than the MDA and ANN models in predicting the bankruptcy risk of the entire industry. These results indicate that the SVM can predict bankrupt companies precisely and is more accurate than the MDA and ANN models in predicting the bankruptcy risk of the entire industry.

Tables 2 to 12 show the result of individual industry prediction. It is conclusion that the results in predicting individual industry bankruptcy indicate that the SVM can predict bankrupt companies precisely, and more accurately than the MDA and ANN models. The result showed that the SVM predicts with 100.00 % accuracy regarding T_p , T_n, and f for 10 industries, except the wholesale industry when 23 financial ratios are used. The result of the wholesale industry showed that T_p is 92.86 % and T_n is 100.00 %, and f is 99.98%. However, when 115 financial ratios are used, the prediction accuracy declines than using 23 financial ratios.

That is, T_p is 100.00 % in the construction, real estate, service, other financing business and information and telecommunication industries. However, T_p is 89.47 % for retail trade, 94.12 % for electrical appliances, 93.75 % for machinery, 71.43 % for wholesale trade, 88.89 % for textile and apparels, and 90.32 % for other industries.

Therefore, we can say that the SVM using 23 financial ratios can predict the bankruptcy risk of companies more precisely than when using 115 financial ratios.

Let's look at the results when we use MDA. In the 23 financial ratios, that T_p equal 100.00 % is no result in any industries, and that T_n equal 100.00 % is only other fining business. In the 115 financial ratios, that T_p equal 100.00 % are real estate and other fining business. That T_n equal 100.00 % are real estate, other fining business, and information and telecommunication. Let's look at the results when we use ANN next. In the 23 financial ratios, that T_p equal 100.00 % is only other fining business. That T_n equal 100.00 % are service, real trade, electric appliances, machinery whole trade, other financing business, textile apparels, and other industry. In the 115 financial ratios, that T_p equal 100.00 % is no result in any industries. That T_n equal 100.00 % are information and telecommunication, and other industry.

The result suggests that the SVM is also more accurate than the MDA and ANN models in predicting the bankruptcy risk for individual industries and the entire industry classification. Furthermore, the precision of prediction for individual industry classification is better than that for the entire industry, and SVM case of using 23 financial ratios is the best model to predict the bankruptcy risk of companies.

The precision and recall for the entire industries are shown in Table 13. The result clearly shows that SVM predicts the bankrupt companies more accurately than MDA and ANN on the precision and recall in case of using 23 financial ratios and 115 financial ratios.

Table 13. Precision and recall for the entire industries (unit:%)

Industry	Methodology	23 financial ratios		115 financial ratios	
		Precision	Recall	Precision	Recall
Entire	MDA	16.41	19.20	15.13	24.11
	ANN	-	0.00	48.34	32.59
	SVM	100.00	99.77	100.00	74.55

The Precision and recall for the individual industry are shown in Table 14. The result clearly shows that SVM predicts the bankrupt companies more accurately than MDA and ANN on the precision and recall, in case of using 23 financial ratios and 115 financial ratios, as the same as the entire industries. Comparing precision and recall of 23 financial ratios and 115 financial ratios, 115 financial ratios showed higher value than 23 financial ratios. By industry, ANN showed high value on the precision and recall in real estate and other financial businesses. This may be related to the number of samples and the number of financial ratios, as described later in the discussion about ANN of Table 16. From now on, a more detailed approach will be needed to solve this reason.

Table 15 shows the results of the learning data (85%), the test data (15%), and the total of these data using ANN. When bankruptcy prediction is performed using the ANN, the search range of the number of neurons in the middle layer and the learning coefficient is pre-set. Learning data (85%) and test data (15%) are randomly generated for all combinations of neuron numbers and learning coefficient, a neural network model is constructed, and the discrimination ratio is calculated. In consideration of the difference in the discrimination accuracy rate by data set, the aforementioned data creation and learning were performed 10 times, and the average value of the discrimination rate was calculated. For example, we would like to show the result of the other financing business using 23 financial ratios. The empirical results from Table 1 to Table 12 using 23 financial ratios and 115 financial ratios are shown; specifically, the results that are the

total of the results of the learning data (85%) and the test data (15%). Therefore, the value of Table 9 using 23 financial ratios is the same as Table 15; the other industry is also similarly exhibited.

Table 14. Precision and recall for individual industry (unit:%)

Industry	Methodology	23 financial ratios		115 financial ratios	
		Precision	Recall	Precision	Recall
Construction	MDA	26.83	25.58	40.00	37.21
	ANN	68.75	25.58	91.18	72.09
	SVM	100.00	100.00	100.00	100.00
Real estate	MDA	90.32	82.35	100.00	100.00
	ANN	96.97	94.12	97.06	97.06
	SVM	100.00	100.00	100.00	100.00
Services	MDA	47.37	45.00	78.57	55.00
	ANN	100.00	55.00	100.00	85.00
	SVM	100.00	100.00	100.00	100.00
Retail trade	MDA	16.00	21.05	26.92	36.84
	ANN	-	0.00	77.78	73.68
	SVM	100.00	100.00	100.00	89.47
Electric appliances	MDA	17.39	23.53	36.36	23.53
	ANN	-	0.00	60.00	0.03
	SVM	100.00	100.00	100.00	94.12
Machinery	MDA	14.29	37.50	28.00	43.75
	ANN	-	0.00	75.00	75.00
	SVM	100.00	100.00	100.00	93.75
Wholesale trade	MDA	15.63	35.71	25.00	50.00
	ANN	-	0.00	88.89	57.14
	SVM	100.00	92.86	100.00	71.45
Other financing business	MDA	100.00	35.71	100.00	100.00
	ANN	100.00	100.00	100.00	92.31
	SVM	100.00	100.00	100.00	100.00
Textile and apparels	MDA	6.67	11.11	21.74	55.56
	ANN	-	0.00	87.50	77.78
	SVM	100.00	100.00	100.00	88.89
Information and telecommunication	MDA	38.46	62.50	100.00	75.00
	ANN	57.14	50.00	-	0.00
	SVM	100.00	100.00	100.00	100.00
Other	MDA	8.51	25.81	9.09	29.03
	ANN	-	0.00	-	0.00
	SVM	100.00	100.00	100.00	90.32

Table 15. Other financing business using 23 financial ratios

Condition		Learning data (85%)				Test data (15%)				Total			
		Status		Total	%	Status		Total	%	Status		Total	%
		0	1			0	1			0	1		
Status	0	12	0	12	100.00	1	0	1	100.00	13	0	13	100.00
	1	0	434	434		0	78	78		0	512	512	
%	0	100.00	0.00	100.00		100.00	0.00	100.00		100.00	0.00	100.00	
	1	0.00	100.00	100.00		0.00	100.00	100.00		0.00	100.00	100.00	

Table 16 shows the results of T_p and T_n using 23 financial ratios and 115 financial ratios in the ANN. It suggests that T_p in the real the estate industry and the other financing business can accurately predict bankruptcy in either the case of 23 financial ratios or 115 financial ratios. One reason for this is that the number of samples in both industries is small,

compared to the number of samples in other industries. In other industries, the number of samples is large, ranging from 2,419 to 25,945. On the other hand, the number of samples is 761 in the real estate industry (34 for bankrupt companies and 727 for non-bankrupt companies), and 525 in the other financing business (13 for bankrupt companies and 512 for non-bankrupt companies). As a result, it can be said that the ANN can predict bankruptcy accurately in cases with small numbers of samples. Conversely, the prediction accuracy of the ANN diminishes in cases with large numbers of samples in this study. However, based on the results of this study, it may not be judicious to come to the above conclusion. In other words, rather than suggest a shortcoming of the ANN, it may actually suggest that the balance between the number of samples and the number of financial indicators used is important. Furthermore, the value of F_n in all the industries shows that 23 financial ratios and 115 financial ratios can accurately predict non-bankrupt companies. On the other hand, characteristic results were obtained in six industries: retail trade, electrical appliances, machinery, wholesale trade, textile and apparels, and other industries. That is, T_p using 23 financial ratios is 0%. However, T_n in both 23 and 115 financial ratios was almost 100.00%. This result shows that ANN can predict non-bankable companies accurately, but can't predict bankrupt companies accurately. To predict accurately the bankruptcy companies needs further refinement.

Table 16. Comparison of 23 financial ratios and 115 financial ratios using test data in ANN (unit: %)

Industry	23 financial ratios		115 indices ratios	
	T_p	T_n	T_p	T_n
Construction	25.58	99.87	72.09	99.92
Real estate	94.12	99.86	97.06	99.86
Services	55.00	100.00	85.00	100.00
Retail trade	0.00	100.00	73.68	99.92
Electric appliances	0.00	100.00	0.03	97.97
Machinery	0.00	100.00	75.00	99.93
Wholesale trade	0.00	100.00	57.14	99.98
Other financing business	100.00	100.00	92.31	100.00
Textile and apparels	0.00	100.00	77.78	99.96
Information and telecommunication	50.00	99.93	0.00	100.00
Other	0.00	100.00	90.32	100.00

Table 17 shows that the results of default kernels for the construction industry using 115 financial ratios as example. The SVM prediction of bankruptcy was conducted using the default kernel in the open source statistical software R. SVM predicted bankruptcy almost perfectly for the entire industries as well as for individual industries. In order to investigate the generalization ability of the model, we use the Gaussian radial basis kernel function and the Laplace radial basis kernel function; the cross was set to 5 and 10, penalty C was set to 10, and a 10-fold cross validation was performed. The default kernel in the open source statistical software R was kernel = rbfdot, kernel = polydot and kernel = laplacedot. As example, we use the various combinations of the kernel function, cross, and penalty. Table 17 shows the result of the combination of kernel = rbfdot, C = 1, cross = 5, the combination of kernel = rbfdot, C=10, and cross=5, and the combination of kernel = laplacedot, C = 10, and cross = 5 in the construction industry using 115 financial ratios. The combination of kernel = laplacedot, C = 10, and cross = 5 was almost perfectly predictable, even in industries other than the other combination, and this combination predicts bankruptcy risk with nearly perfect accuracy in entire industries and individual industries. To sum it up, it turns out that this combination of kernel, C and cross is the best. Thus, the results of the empirical analyses in Table 1 to Table 12 using both the 23 and 115 financial ratios show only this combination.

Table 17. Results of default kernels for the construction industry using 115 financial ratios

Condition		Kernel = rbfdot, C = 1, cross = 5				Kernel = rbfdot, C = 10, cross = 5				Kernel = laplacedot, C = 10, cross = 5			
		Status		Total	%	Status		Total	%	Status		Total	%
		0	1			0	1			0	1		
Status	0	30	13	43	99.67	36	7	43	99.82	43	0	43	100.00
	1	0	3890	3890		0	3890	3890		0	3890	3890	
%	0	69.77	30.23	100.00		83.72	16.28	100.00		100.00	0.00	100.00	
	1	0.00	100.00	100.00		0.00	100.00	100.00		0.00	100.00	100.00	

5. Conclusion

This study predicted the bankruptcy risk of companies listed in the Japanese stock markets based on the entire industry and on individual industries using the MDA, ANN, and SVM, and compared the methods to determine the best method

for predicting bankrupt companies. Initially, this study predicted the bankruptcy of all industries using the three methods without separating the entire industry into individual industries. However, it was impossible to predict bankruptcies accurately with the MDA and ANN. Therefore, we decided to classify the entire industry into 11 industries. Three hypotheses as basic ideas to advance research were initially proposed, and took research methods to verify them. The first hypothesis is adopted when the MDA and ANN were used; however, the hypothesis is rejected in the SVM. That is, the MDA and ANN can predict precisely using many financial ratios, but SVM can predict precisely using fewer financial ratios. Similarly, the second hypothesis is rejected; that is, predicting the bankruptcy risk for one industry type as the entire industry without classification does not result in better accuracy than when it is done with classification. The third hypothesis is adopted by SVM: SVM can predict more accurately than the MDA and ANN.

We conclude from the results of empirical analysis that the SVM is more accurate than the other models in predicting the bankruptcy risk of companies. In the MDA and ANN analysis, bankruptcy prediction could be made accurately only for some individual industries. In contrast, the SVM could predict bankruptcy in companies almost perfectly for either the entire industry and individual industries. This bankruptcy prediction system will assist stakeholders to improve business situations by paying attention to these financial ratios. In future research, we would like to compare additional methods such as Xgboost and logistic regression to determine which is the best method to predict bankrupt companies.

Acknowledgements

The authors would like to thank the editor-in-chief, Prof. Jose Arturo Garza-Reyes and three anonymous reviewers for their insightful comments and suggestions.

References

Altman, E. I., Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, The Journal of Finance, vol. 23, no. 4, pp. 589–609,1968.

Altman, E. I., Marco, G., and Varetto, F., Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience), Journal of banking and finance, vol. 18, no. 3, pp. 505–529, 1994.

Back, B., Laitinen, T., and Sere, K., Neural networks and genetic algorithms for bankruptcy predictions, Expert Systems with Applications, vol. 11, no. 4, pp. 407–413, 1996.

Barniv R., Agarwal, A., and Leach, R., Predicting the outcome following bankruptcy filing-A three-state classification using neural networks, Intelligent Systems in Accounting, Finance and Management, Vol. 6, pp. 177–194, 1997.

Beaver, W. H., Financial ratios as predictors of Failure, Journal of Accounting Research, Empirical Research Accounting: Selected Studies: Supplement, vol. 4, pp. 71–111, 1966.

Bell, T., Neural nets or the logit model? A comparison of each model's ability to predict commercial bank failures, Intelligent Systems in Accounting, Finance and Management, vol. 6, pp. 249–264, 1997.

Boritz, J. E., and Kennedy, D. B., Effectiveness of neural network types for prediction of business failure, Expert Systems with Applications, vol. 9, no. 4, pp. 503–512, 1995.

Boyacioglu, M.A., Kara, Y. and O.K. Baykan, 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, vol.36, pp.3355–3366, 2009

Chang, C. C., Lin, C. J., LIBSVM, A Library for Support Vector Machines, Available: www.csie.ntu. edu.tw/~cjlin / papers/libsvm.pdf. pp. 1–39 (2001, 2013).

Chen, H.Y., A hybrid model for business failure prediction-utilization of particle swarm optimization and support vector machines, Neural Network World, vol. 21, no. 2, pp. 129–152, 2011.

Chen, M.Y. Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches. Computers & Mathematics with Applications, 62, 4514–4524, 2011.

Cheng, C. B., Chen, C. L., and Fu, C. J., Financial distress prediction by a radial basis function network with logit analysis learning. Computers and Mathematics with Applications, vol. 51, pp. 579–588, 2006.

Chung, K. C., Tan, S. S., and Holdsworth, D. K., Insolvency prediction model using multivariate discriminant analysis and artificial neural network for the finance industry in New Zealand, International Journal of Business and Management, vol. 39, no. 1, pp. 19–28, 2008.

Cortes, C., and Vapnik, V., Support-vector networks, Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

Crammer, K., and Singer, Y., On the algorithmic implementation of multiclass kernel-based vector machines, Journal of Machine Learning Research, vol. 2, pp. 265–292, 2001.

Deakin, E. B., A discriminant analysis of predictors of business failure, Journal of Accounting Research, vol. 10, no. 1, pp. 167–179, Spring 1972.

Dong, Y., Hao, X., and Sato, H., Investigation of the impact of data comparability on performance of support vector machine models for credit scoring, Innovation and Supply Chain Management, vol. 9, no. 1, pp. 31–38, 2015.

Fatih Ecer, Comparing the bank failure prediction performance of neural networks, support vector machines: The Turkish case, Ekonomska Istrazivanja-Economic Research, vol.26, no.3, pp.81-98, 2013.

Holland, J. H., Adaptation in natural and artificial systems, University of Michigan Press, Ann Arbor, 1975.

Hsu, C. W., and Lin, C. J., A simple decomposition method for support vector machines, Machine Learning, vol. 46, no. 1, pp. 291–314, 2002.

Montri, I., Veera, B., and Sarun, I., Artificial neural network and genetic algorithm hybrid intelligence for predicting Thai stock price index trend, Computational Intelligence and Neuroscience, vol. 2016, Article ID 3045254, 2016.

Iturriaga, F. J. L., and Sanz, I. P., Bankruptcy visualization and prediction using neural networks: A study of US commercial banks. Expert Systems with Applications, vol. 42, no. 6, pp. 2857–2869, 2015.

Jeong, C., Min, J. H., and Kim, M. S., A tuning method for the architecture of neural network models incorporating GAM and GA as applied to bankruptcy prediction, Expert Systems with Applications, vol. 39, no. 3, pp. 3650–3658, 2012.

Jin, M.(editing), DONG, Y., Data Science of Management and Credit Risk (in Japanese), (Kyoritsu Shuppan Co., Ltd.), pp. 152-186, 2015.

Jo, H., and Han, I., Integration of case-based forecasting, neural network, and discriminant analysis for bankruptcy prediction, Expert Systems with applications, vol. 11, no. 4, pp. 415–422,1996.

Jo, H., Han, I., and Lee, H., Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis. Expert Systems with Applications, 13(2), pp. 97–108, 1997.

Karatzoglou, A., Smola, A., Hornik, K., Zeileis, A.: kernlab – An S4 Package for Kernel Methods in R, Journal of Statistical Software, vol. 11, no. 9, pp. 1–20, 2004.

Kasgari, A. A., Divsalar, M., Javid, M. R., and Ebrahimian, S. J., Prediction of bankruptcy Iranian corporations through artificial neural network and Probit-based analyses, Neural Computing and Applications, vol. 23, pp. 927–936, 2013.

Kim, K. J., Financial time series forecasting using support vector machines, Neurocomputing, vol. 55, no. 1, pp. 307–319, 2003.

Kim, S. Y., Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression, and multivariate discriminant analysis, The Service Industries Journal, vol. 31, no. 3, pp. 441–468, 2011.

Koshika, M., Mastsumaru, M., Kawanaka, T., and Kaneko, S., Bankruptcy prediction model using genetic algorithm and multivariate discriminant analysis, JAMS/JAIMS International Conference on Business & Information 2014 (ICBI2014), Hawaii, USA, September 2–4, 2014.

Lee, S., and Choi, W. S., A multi-industry bankruptcy prediction model using back-propagation neural network and multivariate discriminant analysis, Expert Systems with Applications, vol. 40, no. 8, pp. 2941–2946, 2013.

Matsumaru, M., Kawanaka, T., Katagiri, H., and Kaneko, S., Prediction of bankruptcy on industry classification, International Journal of Japan Association for Management Systems, vol. 10, no. 1, pp. 1–12, 2018.

Min, J. H., and Lee, Y. C., Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters, Expert Systems with Applications, vol. 28, no. 4, pp. 603–614, 2005.

Min, S. H., Lee, J., and Han, I., Hybrid genetic algorithms and support vector machines for bankruptcy prediction, Expert Systems with Applications, vol. 31, no. 3, pp. 652–660, 2006.

Mohatab Rafiei, F., Manzari, S. M., and Bostanian, S., Financial health prediction models using artificial neural networks, genetic algorithm and multivariate discriminant analysis-Iranian evidence, Expert Systems with Applications, vol. 38, pp. 10210–20217, 2011.

Odom, M., and Sharda, R., A neural networks model for bankruptcy prediction, In Proceedings of the IEEE international conference on neural network, vol. 2, 163–168, 1990.

Ohlson, J. A., Financial ratios and the probabilistic prediction of bankruptcy, Journal of Accounting Research, vol. 18, pp. 109–131,1980.

Shin, K. S., Lee, T. S., and Kim, H. J., An application of support vector machines in bankruptcy prediction model, Expert Systems with Applications, vol. 28, no. 1, pp. 127–135, 2005.

Tseng, F. M., and Hu, Y. C., Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks, Expert Systems with Applications, vol. 37, no. 3, pp. 1846–1853, 2010.

Vapnik, V., in S. Haykin (Ed.), Statistical Learning Theory, Adaptive and Learning Systems, vol. 736, John Wiley & Sons Inc, 1998.

Weston, J., Gammerman, A., Stitson, M., Vapnik, V., Vovk, V., and Watkins, C., Support Vector Density Estimation, In B. Scholkopf, C. Burges and A. Smola, Advances in Kernel Methods: Support Vector Machines, MIT Press, Cambridge, MA, pp. 293–306, 1999.

Wilson, R. L., and Sharda, R., Bankruptcy prediction using neural networks, Decision Support Systems, vol. 11, no. 5, pp. 545–557,1994.

Yang, Z., You, W., and Ji, G., Using partial least squares and support vector machines for bankruptcy prediction, Expert Systems with Applications, vol. 38, no. 7, pp. 8336–8342, 2011.

Zhang, G., Hu, Y. M., Patuwo, E. B., and Indro, C. D., Artificial neural networks in bankruptcy prediction, General framework and cross-validation analysis, European Journal of Operational Research, vol. 116, pp. 16–32, 1999.

Zurada, J. M., Malinowski, A., and Cloete, I., Sensitivity analysis for minimization of input data dimension for feedforward neural network. In IEEE International Symposium on Circuits and Systems, ISCAS'94, London, United Kingdom, vol. 6, pp. 447–450,1994.

Appendix A

Table A. 115 financial ratios (1)

Financial index	Variable	Financial index	Variable
Added value	x_1	Management capital	x_2
Working capital 1	x_3	Working capital 2	x_4
Business profit	x_5	Current assets	x_6
Interest-bearing debt	x_7	Current revenue	x_8
Current expenditure	x_9	Real cash flow	x_{10}
Capacity asset	x_{11}	Earnings before interest and taxes	x_{12}
Earnings before interest and taxes, depreciation and amortization	x_{13}	Total capital operating profit margin	x_{14}
Total capital ordinary income ratio	x_{15}	Return on asset (ROA)	x_{16}
Return on equity (ROE)	x_{17}	Ordinary profit ratio of equity capital	x_{18}
Operating capital operating profit margin	x_{19}	Total capital project profit margin	x_{20}
Total capital non-operating income	x_{21}	Total capital internal reserve margin	x_{22}
Interest-bearing debt average interest rate	x_{23}	Ratio of gross profits to sales	x_{24}
Ratio of operating profit to sales	x_{25}	Ratio of ordinary profit to assets	x_{26}
Ratio of pre-tax profit to sales	x_{27}	Ratio of net income to sale	x_{28}
Ratio of non-operating income to sales	x_{29}	Ratio of depreciation to sales	x_{30}
Interest-bearing debt monthly sales ratio	x_{31}	Current account ratio	x_{32}
Finance cost ratio	x_{33}	Ratio of operating expense to sales	x_{34}
Ratio of financial account balance to sales	x_{35}	Ratio of non-operating expenses to sales	x_{36}
Total asset turnover period	x_{37}	Current asset turnover period	x_{38}
Fixed asset turnover period	x_{39}	Tangible fixed asset turnover period	x_{40}
Total assets working capital ratio 1	x_{41}	Total assets working capital ratio 2	x_{42}
Purchase turnover period	x_{43}	Sales debt turnover period	x_{44}
Inventory assets turnover period 1	x_{45}	Inventory assets turnover period 2	x_{46}
Commodity turnover period	x_{47}	Ratio of total assets Interest-bearing debt	x_{48}
Current ratio	x_{49}	Quick ratio	x_{50}
Accounts receivable-to-purchase debt	x_{51}	Defensive interval	x_{52}
Cache interval	x_{53}	Debt ratio	x_{54}
Interest-bearing debt ratio	x_{55}	Interest coverage	x_{56}
Ratio of fixed assets to long-term capital	x_{57}	Fixed ratio	x_{58}
Interest-bearing debt increase rate	x_{59}	Borrowing dependence	x_{60}
Debt rotation period	x_{61}	Fixed liability rotation period	x_{62}
Current liability rotation period	x_{63}	Capital adequacy ratio	x_{64}

Table A. 115 financial ratios (2)

Financial leverage	x_{65}	Short-term borrowing rotation period	x_{66}
Cash deposit monthly commerce ratio	x_{67}	Cash on hand deposit	x_{68}
Debt repayment years 1	x_{69}	Debt return years 2	x_{70}

Debt return years 3	x_{71}	Long-term debt repayment years	x_{72}
Loan Month quotient ratio	x_{73}	Deb capacity ratio	x_{74}
Cash Conversion Cycle 1	x_{75}	Cash conversion cycle 2	x_{76}
Interest rate	x_{77}	Labor productivity 1	x_{78}
Labor productivity 2	x_{79}	Capital investment efficiency 1	x_{80}
Capital investment efficiency 2	x_{81}	Labor distribution rate	x_{82}
Sales to selling, general and administrative	x_{83}	Sales-to-labor cost ratio	x_{84}
Value added rate	x_{85}	Labor equipment ratio 1	x_{86}
Labor equipment ratio 2	x_{87}	Labor equipment ratio 3	x_{88}
Labor equipment ratio 4	x_{89}	Capital productivity 1	x_{90}
Capital productivity 2	x_{91}	Equity share	x_{92}
Borrowed capital share	x_{93}	Sales per person 1	x_{94}
Sales per person 2	x_{95}	Gross profit per person 1	x_{96}
Gross profit per person 2	x_{97}	Operating profit per person 1	x_{98}
Operating profit per person 2	x_{99}	Labor cost per person 1	x_{100}
Labor cost per person 2	x_{101}	Selling, general and administrative	x_{102}
Selling, general and administrative	x_{103}	Total capital growth ratio	x_{104}
Sales growth ratio	x_{107}	Gross profit growth ratio	x_{108}
Operating profit growth ratio	x_{109}	Ordinary profit growth ratio	x_{110}
Pre-tax net income growth ratio	x_{111}	Net income growth ratio	x_{112}
EBIT growth ratio	x_{113}	EBITDA growth ratio	x_{114}
Long and short investment ratio	x_{115}		

Appendix B

Table B. 23 financial ratios

Financial index	Variable	Financial index	Variable
Debt capacity ratio	x_1	Short-term debt rotation period	x_2
Net sales margin ratio	x_3	Debt ratio for total assets	x_4
Loan to value	x_5	Ratio of profit before tax to sales	x_6
Capital share	x_7	Total capital net income margin	x_8
Quick ratio	x_9	Current liability turnover	x_{10}
Total capital ordinary income ratio	x_{11}	Interest-bearing debt monthly sales ratio	x_{12}
Debt monthly sales magnification	x_{13}	Debt turnover period	x_{14}
Ordinary income ratio	x_{15}	Interest coverage ratio	x_{16}
Operating profit per employee	x_{17}	Operating profit per employee	x_{18}
Sales financial balance ratio	x_{19}	Long-term debt repayment years	x_{20}
Finance cost ratio	x_{21}	Ratio of operating profit to operating capital	x_{22}
Return on asset (ROA)	x_{23}		

Table 2. The result of construction

	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
MDA	Status	0	11	32	43	98.42	16	27	43	98.70
		1	30	3860	3890		24	3866	3890	
	%	0	25.58	74.42	100.00		37.21	62.79	100.00	
		1	0.77	99.23	100.00		0.62	99.38	100.00	
ANN	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		

	Status	0	11	32	43	99.06	31	12	43	99.62
		1	5	3885	3890		3	3887	3890	
	%	0	25.58	74.42	100.00		72.09	27.91	100.00	
		1	0.13	99.87	100.00		0.08	99.92	100.00	
SVM	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
	Status	0	43	0	43	100.00	43	0	43	100.00
		1	0	3890	3890		0	3890	3890	
	%	0	100.00	0.00	100.00		100.00	0.00	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	

Table 3. The result of real estate

MDA	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
	Status	0	28	6	34	98.82	34	0	34	100.00
		1	3	724	727		0	727	727	
	%	0	82.35	17.65	100.00		100.00	0.00	100.00	
		1	0.41	99.59	100.00		0.00	100.00	100.00	
ANN	Item		23 financial ratios				115 financial ratios			
			Status		Total	**%**	Status		Total	**%**
			0	1			0	1		
	Status	0	32	2	34	99.61	33	1	34	99.74
		1	1	726	727		1	726	727	
	%	0	94.12	5.88	100.00		97.06	2.94	100.00	
		1	0.14	99.86	100.00		0.14	99.86	100.00	
SVM	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
	Status	0	34	0	34	100.00	34	0	34	100.00
		1	0	727	727		0	727	727	
	%	0	100.00	0.00	100.00		100.00	0.00	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	

Table 4. The result of services

MDA	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
	Status	0	9	11	20	99.48	11	9	20	99.71
		1	10	4047	4057		3	4054	4057	
	%	0	45.00	55.00	100.00		55.00	45.00	100.00	
		1	0.25	99.75	100.00		0.07	99.93	100.00	
ANN	Item		23 financial ratios				115 financial ratios			
			Status		Total	**%**	Status		Total	**%**
			0	1			0	1		
	Status	0	11	9	20	99.78	17	3	20	99.93
		1	0	4057	4057		0	4057	4057	
	%	0	55.00	45.00	100.00		85.00	15.00	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	
SVM	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
	Status	0	20	0	20	100.00	20	0	20	100.00
		1	0	4057	4057		0	4057	4057	
	%	0	100.00	0.00	100.00		100.00	0.00	100.00	

		1	0.00	100.00	100.00		0.00	100.00	100.00	

Table 5. The result of retail trade

	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
MDA	Status	0	4	15	19	99.28	7	12	19	99.38
		1	21	4947	4968		19	4949	4968	
	%	0	21.05	78.95	100.00		36.84	63.16	100.00	
		1	0.42	99.58	100.00		0.38	99.62	100.00	
ANN	Item		23 financial ratios				115 financial ratios			
			Status		Total	**%**	Status		Total	**%**
			0	1			0	1		
	Status	0	0	19	19	99.62	14	5	19	99.82
		1	0	4968	4968		4	4964	4968	
	%	0	0.00	100.00	100.00		73.68	26.32	100.00	
		1	0.00	100.00	100.00		0.08	99.92	100.00	
SVM	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
	Status	0	19	0	19	100.00	17	2	19	99.96
		1	0	4968	4968		0	4968	4968	
	%	0	100.00	0.00	100.00		89.47	10.53	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	

Table 6. The result of electric appliances

	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
MDA	Status	0	4	13	17	99.50	4	13	17	99.69
		1	19	6407	6426		7	6419	6426	
	%	0	23.53	76.47	100.00		23.53	76.47	100.00	
		1	0.30	99.70	100.00		0.11	99.89	100.00	
ANN	Item		23 financial ratios				115 financial ratios			
			Status		Total	**%**	Status		Total	**%**
			0	1			0	1		
	Status	0	0	17	17	99.74	3	14	17	99.75
		1	0	6426	6426		2	6424	6426	
	%	0	0.00	100.00	100.00		17.65	82.35	100.00	
		1	0.00	100.00	100.00		0.03	99.97	100.00	
SVM	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
	Status	0	17	0	17	100.00	16	1	17	99.98
		1	0	6426	6426		0	6426	6426	
	%	0	100.00	0.00	100.00		94.12	5.88	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	

Table 7. The result of machine

	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
MDA	Status	0	6	10	16	99.15	7	9	16	99.50
		1	36	5369	5405		18	5387	5405	
	%	0	37.50	62.50	100.00		43.75	56.25	100.00	
		1	0.67	99.33	100.00		0.33	99.67	100.00	

	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
ANN	Status	0	0	16	16	99.70	12	4	16	99.85
		1	0	5405	5405		4	5401	5405	
	%	0	0.00	100.00	100.00		75.00	25.00	100.00	
		1	0.00	100.00	100.00		0.07	99.93	100.00	
SVM	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
	Status	0	16	0	16	100.00	15	1	16	99.98
		1	0	5405	5405		0	5405	5405	
	%	0	100.00	0.00	100.00		93.75	6.25	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	

Table 8. The result of wholesale trade

	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
MDA	Status	0	5	9	14	99.40	7	7	14	99.53
		1	27	5944	5971		21	5950	5971	
	%	0	35.71	64.29	100.00		50.00	50.00	100.00	
		1	0.45	99.55	100.00		0.35	99.65	100.00	
ANN	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
	Status	0	0	14	14	99.77	8	6	14	99.88
		1	0	5971	5971		1	5970	5971	
	%	0	0.00	100.00	100.00		57.14	42.86	100.00	
		1	0.00	100.00	100.00		0.02	99.98	100.00	
SVM	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
	Status	0	13	1	14	99.98	10	4	14	99.93
		1	0	5971	5971		0	5971	5971	
	%	0	92.86	7.14	100.00		71.43	28.57	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	

Table 9. The result of other financing business

	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
MDA	Status	0	12	1	13	99.81	13	0	13	100.00
		1	0	512	512		0	512	512	
	%	0	92.31	7.69	100.00		100.00	0.00	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	
ANN	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
	Status	0	13	0	13	100.00	12	1	13	99.81
		1	0	512	512		0	512	512	
	%	0	100.00	0.00	100.00		92.31	7.69	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	
SVM	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		

	Status	0	13	0	13	100.00	13	0	13	100.00
		1	0	512	512		0	512	512	
	%	0	100.00	0.00	100.00		100.00	0.00	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	

Table 10. The result of textile and apparels

	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
MDA	Status	0	1	8	9	99.09	5	4	9	99.09
		1	14	2396	2410		18	2392	2410	
	%	0	11.11	88.89	100.00		55.56	44.44	100.00	
		1	0.58	99.42	100.00		0.75	99.25	100.00	
	Item		23 financial ratios				115 financial ratios			
			Status		Total	**%**	Status		Total	**%**
			0	1			0	1		
ANN	Status	0	0	9	9	99.63	7	2	9	99.88
		1	0	2410	2410		1	2409	2410	
	%	0	0.00	100.00	100.00		77.78	22.22	100.00	
		1	0.00	100.00	100.00		0.04	99.96	100.00	
	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
SVM	Status	0	9	0	9	100.00	8	1	9	99.96
		1	0	2410	2410		0	2410	2410	
	%	0	100.00	0.00	100.00		88.89	11.11	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	

Table 11. The result of Information and telecommunication

	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
MDA	Status	0	5	3	8	99.75	6	2	8	99.95
		1	8	4420	4428		0	4428	4428	
	%	0	62.50	37.50	100.00		75.00	25.00	100.00	
		1	0.18	99.82	100.00		0.00	100.00	100.00	
	Item		23 financial ratios				115 financial ratios			
			Status		Total	**%**	Status		Total	**%**
			0	1			0	1		
ANN	Status	0	4	4	8	99.84	0	8	8	99.82
		1	3	4425	4428		0	4428	4428	
	%	0	50.00	50.00	100.00		0.00	100.00	100.00	
		1	0.07	99.93	100.00		0.00	100.00	100.00	
	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
SVM	Status	0	8	0	8	100.00	8	0	8	100.00
		1	0	4428	4428		0	4428	4428	
	%	0	100.00	0.00	100.00		100.00	0.00	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	

Table 12. The result of other industries

	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
MDA	Status	0	8	23	31	99.58	9	22	31	99.57

		1	86	25828	25914		90	25824	25914	
	%	0	25.81	74.19	100.00		29.03	70.97	100.00	
		1	0.33	99.67	100.00		0.35	99.65	100.00	
ANN	Item		23 financial ratios				115 financial ratios			
			Status		Total	**%**	Status		Total	**%**
			0	1			0	1		
	Status	0	0	31	31	99.88	0	31	31	99.88
		1	0	25914	25914		0	25914	25914	
	%	0	0.00	100.00	100.00		0.00	100.00	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	
SVM	Item		23 financial ratios				115 financial ratios			
			Status		Total	%	Status		Total	%
			0	1			0	1		
	Status	0	31	0	31	100.00	28	3	31	99.99
		1	0	25914	25914		0	25914	25914	
	%	0	100.00	0.00	100.00		90.32	9.68	100.00	
		1	0.00	100.00	100.00		0.00	100.00	100.00	

Biographies

Masanobu Matsumaru is currently a visiting professor in Research Institute for Engineering at Kanagawa University, and was a former professor of Department of Industrial Engineering and Management at Kanagawa University, Japan. He earned Bachelor of Engineering in Faculty of Science and Engineering from Waseda University, Japan, Master of Engineering in Master course Graduate school of Science and Engineering from Waseda University and Doctor of Engineering from Tokai University. His research interests include Industrial Engineering and Management, Management Modeling, Problem Solving, Financial Management and Supply Chain.

Takaaki Kawanaka is currently a lecturer in Institute for Innovation in International Engineering Education, Graduate School of Engineering, the University of Tokyo, Japan. He earned Bachelor Degree of Engineering in Faculty of Science and Engineering from Waseda University, Japan, Master Degree of Engineering in Master course Graduate school of Science and Engineering from Waseda University and Doctor Degree of Engineering in Doctor course Graduate school of Engineering from the University of Tokyo, Japan. His research interests are mainly focused on Industrial Engineering, Management Modeling and Information Security Management.

Shoichi Kaneko is a Professor of faculty of Management Information at Yamanashi Gakuin University, in Yamanashi, JAPAN. He earned B.E. in Faculty of Science and Engineering and earned M.E. in Master Course Graduate School of Science from Waseda University. He engaged in precision machinery company for work such as Total Quality Management, Management Planning, etc. He is studying Management Quality Science, Management Model, Human Resources Management and so on. He is member of JAMS(Japan Association for Management Systems) and JIMA(Japan Industrial Management Association).

Hideki Katagiri is a Professor of Department of Industrial Engineering and Management at Kanagawa University, Japan. He is a Visiting Professor of Hiroshima University, Japan. He earned his B.E., M.E. and Ph.D. in Engineering at Osaka University in 1995, 1997 and 2000, respectively. He was the Chair of IEEE SMC Hiroshima Section Chapter (2008-2010) and a Visiting Scholar at the University of Chicago Booth School of Business (2014-2015). His research and teaching activities are in the areas of operations research and soft computing, especially, multi-objective optimization under uncertainty and data analysis using machine learning techniques. He is the author or co-author of more than 100 refereed journal papers and several co-authored or co-edited books in English.