World Academy of Science, Engineering and Technology
International Journal of Economics and Management Engineering
Vol:5, No:12, 2011

Corporate Credit Rating using Multiclass Classification Models with order Information

Hyunchul Ahn and Kyoung-Jae Kim

***Abstract*—**Corporate credit rating prediction using statistical and artificial intelligence (AI) techniques has been one of the attractive research topics in the literature. In recent years, multiclass classification models such as artificial neural network (ANN) or multiclass support vector machine (MSVM) have become a very appealing machine learning approaches due to their good performance. However, most of them have only focused on classifying samples into nominal categories, thus the unique characteristic of the credit rating - ordinality - has been seldom considered in their approaches. This study proposes new types of ANN and MSVM classifiers, which are named OMANN and OMSVM respectively. OMANN and OMSVM are designed to extend binary ANN or SVM classifiers by applying ordinal pairwise partitioning (OPP) strategy. These models can handle ordinal multiple classes efficiently and effectively. To validate the usefulness of these two models, we applied them to the real-world bond rating case. We compared the results of our models to those of conventional approaches. The experimental results showed that our proposed models improve classification accuracy in comparison to typical multiclass classification techniques with the reduced computation resource.

***Keywords*—**Artificial neural network, Corporate credit rating, Support vector machines, Ordinal pairwise partitioning

I. INTRODUCTION

CORPORATE credit ratings are very important element in the market for corporate debt. Information concerning corporate operations is often disseminated to market participants through the changes in credit ratings published by professional rating agencies such as Standard & Poor's (S&P) and Moody's Investor Service. Since these agencies generally require a large amount of fee for the service, and the ratings provided periodically do sometimes not reflect the default risk of the company in time, it may be advantageous to the bond market participants to be able to classify credit ratings before the ratings are published by the agencies. As a result, it is very important for companies (especially, financial companies) to develop a proper model for credit rating [1].

From the perspective of the technical aspects, the credit rating is a typical multiclass classification problem because the rating agencies have ten or more categories of ratings. The professional rating agencies emphasize the importance of analysts' subject judgment in determining credit ratings. However, in practice, a mathematical model using financial variables of companies plays an important role in determining credit ratings since it is convenient to apply, and requires less time and cost.

As tools for credit ratings prediction, several statistical and artificial intelligence (AI) techniques have been applied. Among them, artificial neural network (ANN) has been widely used in the area of finance because of their broad applicability to many business problems and preeminent learning ability. Recently, support vector machine (SVM) also becomes popular as a solution for prediction problems because of their robustness and high accuracy. But, SVMs were originally devised for binary classification, so it doesn't fit exactly to multiclass classification just like credit rating [2]. Thus, researchers have tried to extend it to multiclass classification. As a result, various approaches of multiclass SVM (MSVM) are proposed up to now.

However, typical ANN and MSVM have focused on classifying samples into nominal categories [3-8]. Even the prior studies applied ANNs or MSVMs to credit rating also applied the typical models that were not designed to reflect the ordinal characteristic of this domain [1,9-11].In this study, we propose a novel approach for multiclass classification, which takes ordinal characteristics into account in order to handle ordinal multiple classes efficiently and effectively. Our models basically combine several binary ANN or SVM classifiers. However, they are different from traditional approaches since they are designed to extend binary classifiers using ordinal pairwise partitioning (OPP) approach [12]. The application of OPP makes them use less classifier, but may predict more accurately because they exploit additional hidden information, the order of the classes. To empirically validate the usefulness of our models, we apply them to a real-world bond rating case. We compared the results of the models to those of traditional ANN and MSVM approaches.

II. THEORETICAL BACKGROUND

A. Artificial neural network

ANN is a computing technology that captures the salient fundamental features of the inputs and recognizes the pattern in the data. The ANN model can discover previously unknown or complex nonlinear relationships in certain data. The ANN model shares the human brain's capacity to learn from a repeated number of inputs, by adjusting the weights that are assigned to the neurons. Through such training, the ANN model can be utilized for both binary and multiclass classification tasks, such as exchange rate or stock price direction prediction,

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (KRF – 2009 – 332 – B00104).

H. Ahn is with the School of Management Information Systems, Kookmin University, 861-1 Jeongneung-dong, Seongbuk-gu, Seoul 136-702, Republic of Korea (phone: +82-2-910-4577; fax: +82-2-910-5209; e-mail: hcahn@kookmin.ac.kr).

K.-j. Kim is with the Department of Management Information Systems, Dongguk University-Seoul, 3-26 Pil-dong, Chung-gu, Seoul 100-715, Republic of Korea (e-mail: kjkim@dongguk.edu).

credit rating, churn prediction and even box office success predictions [13].

B. Multiclass support vector machine

Conventional SVMs were originally designed for binary classification, which has only one classifier. Thus, in order to perform multi-class classification, conventional SVMs should be modified, and the topics for extending conventional SVMs to MSVMs are still ongoing research issues. In recent five years, a variety of techniques for implementing MSVMs have been proposed. In general, there are two approaches for the extension of SVMs to multiclass problems. The first approach is to decompose the multiclass problems in several binary subproblems. In this approach, MSVMs can be implemented by constructing and combining several binary SVM classifiers. The second approach is directly considering all data in one optimization formulation. In this case, the modification of the conventional training algorithm of SVMs is required. There are different kinds of techniques in each approach [14,15].

In general, there are three popular algorithms for constructing and combining several binary SVM classifiers. The first algorithm is One-Against-All. This is conceptually the simplest multiclass method. This method constructs k binary SVM classifiers for k-class classification: class 1 (positive) versus all other classes (negative), class 2 versus all other classes, … , class k versus all other classes. The second algorithm is One-Against-One. In this method, the model constructs binary SVM classifiers for all pairs of classes; in total there are ${}_kC_2$ pairs. It means, for every pair of classes, a binary SVM classifier is constructed by solving the underlying optimization problem to maximize the margin between two classes. The decision function assigns an instance to a class which has the largest number of votes, so-called *Max Wins* (same as vote count, or winner-takes-all) strategy [16,17]. The third algorithm is the directed acyclic graph SVM (DAGSVM). The training phase of this algorithm is similar to the One-Against-One method using multiple binary classifiers; however DAGSVM uses graph-visiting strategy for testing. The testing phase of DAGSVM requires construction of a rooted binary decision directed acyclic graph (DDAG) using classifiers. Each node of this graph is a binary SVM for a pair of classes, say (p, q). On the topologically lowest level, there are k leaves corresponding to k classification decisions. Every non-leaf node (p, q) has two edges – the left edge corresponds to decision “not p” and the right one corresponds to “not q”. The choice of the class order in the DDAG list can be arbitrary as shown empirically in [4].

As the algorithms for directly considering all data at once, the method by Weston and Watkins, and the method by Crammer and Singer are generally used. These algorithms may be interpreted as a natural extension of the binary SVM classification problem. Here, in the k-class case, one has to solve single quadratic optimization problem of size $(k-l)n$ which is identical to a binary SVM for the case of k=2 [18]. In a slightly different formulation of QP problem, a bounded formulation, decomposition technique can provide a significant speed-up in the solution of the optimization problem [5,19]. The method by Crammer and Singer is similar to the method by Weston and Watkins, however, it uses less slack variables in the constraints of the optimization problem [3].

Among these five MSVM algorithms, One-Against-One is used most popularly because it’s the simplest, but its prediction performance is generally good [9].

C. Credit rating using artificial intelligence techniques

Substantial studies on bond rating prediction using data mining techniques can be found in the literature. These studies can be categorized into three stages. The early days of these studies mainly focused on applicability of statistical techniques such as multiple discriminant analysis (MDA) and logistic regression analysis (LogR). The second stage of the research is application of typical artificial intelligence techniques such as artificial neural networks (ANNs) and case-based reasoning (CBR). Backpropagation neural networks, (BPNs), a kind of ANNs, had been applied most popularly in this stage. However, they suffer from difficulty in selecting a large number of controlling parameters which include relevant input variables, hidden layer size, learning rate, and momentum term. In addition, they require a large amount of data for training model due to degree of freedom constraint. To overcome these limitations, recent studies try to apply MSVMs to credit rating.

The study by Huang et al. [9] is the pioneer of the studies that adopted MSVMs for building prediction model of credit rating. They experimented various techniques of MSVMs including One-Against-One and the method by Crammer and Singer. They also experimented with the different parameters in order to find optimal MSVM model. Finally, they selected the method by Crammer and Singer using a RBF kernel function. They found that this MSVM model outperformed not only BPNs, but also LogR in the prediction of Taiwan and US bond rating.

Cao et al. [1] applied One-Against-All, One-Against-One, and DAGSVM to the prediction of S&P’s bond ratings. For the kernel function, Gaussian RBF is applied and the optimal kernel parameters were investigated by grid search strategy. As a result, they found that DAGSVM showed the best performance among them, and all kinds of MSVM approaches outperformed other comparative multi-class classification technique including LogR, ordered probit regression (OPR) and BPNs.

Chen and Shih [10] adopted One-Against-One approach to build the automatic classification model for issuer credit ratings in Taiwan. Similar to the study by Cao et al. [1], they also adopted Gaussian RBF kernel function and the grid search strategy for searching optimal parameters. They found that MSVM model was superior to BPN and LogR models with the statistical significance.

The study by Lee [11] applied the same approach with Chen and Shih [10] for building a corporate credit rating prediction model for Korean companies. He also adopted Gaussian RBF kernel function and the grid search strategy. The experimental results showed that MSVM model significantly outperformed BPNs, MDA and CBR.

III. Ordinal Multiclass Classification Models

Up to now, the multiclass classification techniques applied in corporate credit ratings are mostly designed for the multiclass classification problems whose classes are nominal or categorical, not ordinal. However, a proper modification of the conventional multiclass classification techniques by considering ordinality may improve the performance because of *information gain* effects.

For this motivation, we propose new types of ANN and MSVM models that are optimized for ordinal multiclass classification problems just like credit rating. Our approaches, named OMANN(Ordinal Multiclass ANN) and OMSVM (Ordinal Multiclass SVM), are hybrid algorithms that apply the concept of ordinal pairwise partitioning (OPP) technique to typical ANN and MSVM.

OPP is the approach that is designed to enhance the performance of artificial neural networks model for ordinal multiclass classification, proposed by Kwon et al. [12]. The authors noticed that ANNs that are designed to predict multiclass classification problems generally show worse performance than combining several binary ANN classifiers. Thus, they proposed a new method called OPP to combine several binary ANN classifiers considering the order.

Thus, in this study, we suggest ordinal pairwise partitioning (OPP) approach as a tool for upgrading conventional ANN as well as MSVM models in order to deal with ordinal classes wisely. The OPP approach partitions the data set into subdata sets with reduced classes in the ordinal and pairwise manner according to the output classes. As shown in Table I, there are four types of OMANNs and OMSVMs according to the partitioning methods and the fusing methods.

For the partitioning method, there are One-Against-TheNext and One-Against-Followers approaches. One-Against-TheNext method is similar to One-Against-One, however, it's much more efficient. In the case of One-Against-One, all the classifiers for each pair of classes should be developed. But, in One-Against-TheNext, the binary classifiers for the pairs $(i, i+1)$ are constructed where $i = 1, 2, \cdots, k-1$ and k is the total number of classes. Consequently, One-Against-TheNext only constructs $k-1$ binary classifiers where there are k classes.

Contrast to One-Against-TheNext, One-Against-Followers is similar to One-Against-All. But, it is also a little bit more efficient than One-Against-All. In the case of One-Against-Followers, the binary classifiers for the pairs (i, j) are constructed where $i = 1, 2, \cdots, k-1$, $j = \bigcup_{m=i+1}^{k} m$, and k is the total number of classes. As a result, One-Against-Followers also only constructs $k-1$ binary classifiers, although One-Against-All constructs k classifiers where there are k classes.

Regarding fusing methods, there are forward and backward methods, which implies the *reasoning direction*. The forward method fuses the binary classifiers in forward direction – that is, it determines the highest level of classes first, and the lowest level last. By contrast, the backward one combines the binary classifiers in reverse direction. Thus, it determines the lowest level of classes first, and the highest level last.The process of OMANN or OMSVM consists of two phases – (1) preparation, and (2) interpretation phase. In preparation phase, the proposed models construct individual binary classifiers using training data set. In detail, it first divides the whole training data set into $k-1$ groups according to the partitioning method. Then, it trains $k-1$ binary ANN or SVM models with each of the divided data sets above. For example, when using One-Against-TheNext approach, the first phase of OMANN or OMSVM produces three binary classification models – Model 1 for the pair of classes (1, 2), Model 2 for the pair of classes (2, 3), and Model 3 for the pair of classes (3, 4) – for 4-level classification problems.

TABLE I
Four Types of Proposed Models and Their Example Process of a 4-class Classification Problem

		Partitioning Method	
		One-Against-TheNext	One-Against-Followers
Fusing Method	For-ward	*Trains the following classifiers: (1vs2), (2vs3), (3vs4)* *1)* Apply the classifier (1vs2) → Determine the class 1 *2)* Apply the classifier (2vs3) → Determine the class 2 *3)* Apply the classifier (3vs4) → Determine the class 3 and 4	*Trains the following classifiers: (1vs2,3,4), (2vs3,4), (3vs4)* *1)* Apply the classifier (1vs2,3,4) → Determine the class 1 *2)* Apply the classifier (2vs3,4) → Determine the class 2 *3)* Apply the classifier (3vs4) → Determine the class 3 and 4
	Back-ward	*Trains the following classifiers: (1vs2), (2vs3), (3vs4)* *1)* Apply the classifier (3vs4) → Determine the class 4 *2)* Apply the classifier (2vs3) → Determine the class 3 *3)* Apply the classifier (1vs2) → Determine the class 1 and 2	*Trains the following classifiers: (1vs2), (1,2vs3), (1,2,3vs4)* *1)* Apply the classifier (1,2,3vs4) → Determine the class 4 *2)* Apply the classifier (1,2vs3) → Determine the class 3 *3)* Apply the classifier (1vs2) → Determine the class 1 and 2

In interpretation phase, our proposed models determine the class for the input data using the binary classifiers built in the first phase. To do this, it fuses the binary classifiers in the forward direction or backward direction. In the case of above example, the forward method begins with Model 1. If a test data is put into class 1 by Model 1, it is called class 1. Otherwise, the test data is passed on to Model 2. If it is put into class 2 by Model 2, then it is called class 2. Otherwise, Model 3 applies. In Model 3, the test data is finally classified as either class 3 or class 4. Using the same reasoning, the backward method starts with Model 3. That is, if test data is put into class 4, we regard it as class 4. Otherwise, the test data is passed on to the next ANN or SVM model. The remaining procedure is the same as in the forward method but in the reverse order.

IV. Experimental Design

To validate our models, we applied them to the real world credit rating case in South Korea. Our application is bond rating, which is the most frequently studied credit rating area for specific debt issues or other financial obligations. The research data was collected from National Information and Credit Evaluation, Inc., a major bond rating company in South Korea. We obtained the bond rating results for the year 2002 and various financial variables from 1,295 companies in manufacturing industry in Korea. In Korean bond rating market, bond ratings are divided into 5 classes, A1, A2, A3, B and C. But, we adjust our data to 4 classes by combining B and C ratings into one group because their numbers of samples were so small, and these ratings are usually treated same as just junk bonds in the market.

TABLE IV
The Experimental Results of Typical ANN and MSVM

Comparative Models	Data Set	Hit Ratio
Typical ANN	Train	69.55%
	Test	65.35%
	Validation	65.66%
Typical MSVM (DAGSVM)	Train/Test	70.84%
	Validation	67.29%

Original data consisted of 39 financial ratio variables that were known to the features affecting bond rating in previous literatures. Among them, we selected 36 variables by applying independent-samples *t*-test and, finally, selected 14 variables which are proved to be the most influential in bond rating by applying stepwise statistical method. The selected variables are presented in Table II. In this study, 20% of the data for each class were used for validation and the remaining 80% of data were used for training. And, to overcome the scarcity of samples, we adopted 5-fold cross-validation.

TABLE II
Definition of the Selected Input Variables

Symbol	Definition
SHEQ	Shareholder's equity
SALE	Sales
DEBT	Total debt
SAPE	Sales per employee
NIPS	Net income per share
YEAR	Years after founded
AETA	Accumulated earning to total asset
BDRA	Borrowings-dependency ratio
FCTC	Financing cost to total cost
FIRA	Fixed ratio
IACA	Inventory assets to current assets
SBTB	Short-term borrowings to total borrowings
CFTA	Cash flow to total assets
OACF	Cash flows from operating activity

To validate the superiority of our models' performances with sophistication, we applied our proposed models (OMANN and OMSVM) as well as typical ANN and MSVM. As the typical MSVM, we selected DAGSVM algorithm since a prior study reported that it showed the best prediction accuracy among traditional MSVM algorithms [1].

In the case of MSVM models (OMSVM and DAGSVM), the linear kernel, the polynomial kernel and the Gaussian radial basis function are used as the kernel function of SVM. Tay and Cao [20] showed that the upper bound C and the kernel parameter play an important role in the performance of SVMs. Improper selection of these two parameters can cause the overfitting or the underfitting problems.

Since there is few general guidance to determine the parameters of SVM, this study varies the parameters including upper bound C, d of polynomial kernel function, and σ^2 of Gaussian RBF kernel function, to select optimal values for the best prediction performance. For the implementation of OMSVM and DAGSVM, we developed post-processing software that combines the results in predefined ways generated from multiple binary SVM classifiers. It was written in Microsoft Visual Basic for Applications for Excel 2003. And, as a library for binary SVM classification, we adopted LIBSVM 2.6 provided by Chang and Lin [21].

In the case of ANNs, we adopt standard three-layer back propagation networks and set the number of nodes in the hidden layer as 7, 14, 21 and 28. For the stopping criteria of ANNs, this study allows 50 learning epochs and set the learning rate to 0.1 and the momentum term to 0.1. The hidden nodes use the sigmoid transfer function and the output node uses the linear transfer function. This study allows 14 input nodes because 14 input variables are employed. We implement binary and multiclass ANN models using Neuroshell R4.0. OMANN, similar to OMSVM, is experimented using the post-processing software that combines the results in predefined ways generated from multiple binary ANN classifiers.

V. Experimental Results

A. OMSVM and Typical MSVM (DAGSVM)

Table III shows the hit ratios of the proposed model, OMSVM. As shown this table, One-Against-Followers + Forward approach showed the best performance (67.98%). For the same fusing method, the prediction accuracies of One-Against-Followers approach were always higher than ones of One-Against-TheNext (67.98%>67.36% for forward strategy, 67.60%>67.13%). And, for the same partitioning method, the forward fusing method always outperformed the reverse fusing method (67.36%>67.13% for One-Against-TheNext, 67.98%>67.60% for One-Against-Followers).

Table IV presents the experimental results of the comparative models. From this table, we can check that OMSVM outperformed conventional MSVM (i.e. DAGSVM). In particular, we are also able to find that OMSVM regardless of its types always showed better prediction performance than DAGSVM except for One-Against-TheNext+Backward type.

B. OMANN and Typical ANN

In Table V, the results of OMANN – ANN model that adopted OPP – are presented. OMANN produced the prediction accuracy ranging from 66.43% to 67.05%. When considering the prediction accuracy of conventional ANN for the validation

TABLE III
THE EXPERIMENTAL RESULTS OF OMSVM

Partitioning Method	Fusing Method	Data Set	Hit Ratio
One-Against-TheNext	Forward	Train	71.80%
		Validation	67.36%
	Backward	Train	71.98%
		Validation	67.13%
One-Against-Followers	Forward	Train	75.77%
		Validation	67.98%[a]
	Backward	Train	73.58%
		Validation	67.60%

[a]The best performance for the validation data set.

data set was 65.66% presented in Table IV, it is believed that OPP approach is able to improve the performances of SVMs as well as ANNs. Nevertheless, its best prediction accuracy (67.05%) was lower than even OMSVM's worst result (67.13%). In general, SVMs are known that they produce more accurate. Thus, it is believed that it would be more effective for bond rating to apply OMSVM rather than OMANN.It is also interesting that the partitioning approach that shows the best performance of OMSVM and OMANN is completely different. As indicated in Table III, One-Against-Followers outperformed One-Against-TheNext in OMSVM. However, in the case of OMANN, One-Against-TheNext outperformed One-Against-Followers. In the study of Kwon et al. [12], we can also find the same pattern. We think this phenomenon is caused by the fundamental difference between ANN and SVM. As mentioned before, ANN is the method implements the *Empirical Risk Minimization* principal, and is designed to minimize the training error by repeated learning. Thus, it is easy to be affected by disproportion of the samples. By contrast, SVM is basically free from the problem of sample disproportion because SVM just refers a small subset of training samples, which is called support vectors. In other words, it implements the *Structural Risk Minimization* principal [1]. One-Against-Followers approach may provide more detailed information for classification, however, the proportion of the training sample is distorted. As a result, SVM by nature can fully utilize the advantage of One-Against-Followers approach, but for ANN, misleading the training may happen when using this approach, and it may result in low prediction performance.

TABLE V
THE EXPERIMENTAL RESULTS OF OMANN

Partitioning Method	Fusing Method	Data Set	Hit Ratio
One-Against-TheNext	Forward	Train	68.34%
		Test	66.43%
		Validation	67.05%[a]
	Backward	Train	68.27%
		Test	66.43%
		Validation	66.98%
One-Against-Followers	Forward	Train	67.98%
		Test	66.90%
		Validation	66.36%
	Backward	Train	67.52%
		Test	66.59%
		Validation	66.43%

[a]The best performance for the validation data set.

VI. CONCLUDING REMARKS

In this study, we proposed a novel multiclass classification models optimized for credit rating. Contrast to prior studies that just applied conventional ANNs or MSVMs to credit ratings, we suggested a new multiclass classification models, called OMANN and OMSVM, which are designed to use order information when classifying ordinal multiclass problems. To validate the applicability of the proposed models, we applied them to real bond rating case. As a result, we found that OMSVM outperformed a typical MSVM approach, and OMANN also outperformed a typical ANN. And among them, the performance of OMSVM was better than one of OMANN. As a result, we may conclude that OMSVM is effective and efficient classifier for solving ordinal multiclass classification problems like corporate credit rating.

Although we applied our models to the domain of credit rating here, they can be applied to any kinds of ordinal multiclass classification problems. For example, in medical diagnostics, doctors may want to build prediction model that classifies patients by the level of severity of a disease. And, in business domain, some marketers may want to build classification model that classifies customers by the level of profitability in order to implement customer relationship management (CRM) strategy. Besides, there are many kinds of application areas, which require accurate ordinal multiclass classification model. Thus, we expect that our proposed models will be able to contribute to other domains or business problems in the future studies.

REFERENCES

[1] L. Cao, L. K. Guan, and Z. Jingqing, "Bond rating using support vector machine," *Intell. Data Anal.*, vol. 10, no. 3, pp. 285-296, 2006.
[2] V. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995.
[3] K. Crammer, and Y. Singer, "On the learnability and design of output codes for multiclass problems," in *Proc. 13th Annu. Conf. Computational Learning Theory*, Palo Alto, CA, 2000, pp. 35-46.
[4] J. C. Platt, N. Cristianini, and J. Shawe-Taylor, "Large margin DAG's for multiclass classification," in *Advances in Neural Information Processing Systems*, vol. 12, S. A. Solla, T. K. Leen, and K. -R. Muller, Eds. Cambridge, MA: MIT Press, 2000, pp. 547-553.
[5] C. -W. Hsu, and C. -J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Networks*, vol. 13, no. 2, pp. 415-425, 2002.
[6] Z. Shuibo, T. Houjun, H. Zhengzhi, and Z. Haoran, "Solving large-scale multiclass learning problems via an efficient support vector classifier," *J. Syst. Eng. Electron.*, vol. 17, no. 4, pp. 910-915, 2006.
[7] A. Navia-Vázquez, "Compact multi-class support vector machine," *Neurocomputing*, vol. 71, nos. 1-3, pp. 400-405, 2007.
[8] E. D. Übeyli, "Multiclass support vector machines for diagnosis of erythemato-squamous disease," *Expert Syst. Appl.*, vol. 35, no. 4, pp. 1733-1740, 2008.
[9] Z. Huang, H. Chen, C. -J. Hsu, W. -H. Chen, and S. Wu, "Credit rating analysis with support vector machines and neural networks: A market comparative study," *Decis. Support Syst.*, vol. 37, no. 4, pp. 543-558, 2004.
[10] W. -H. Chen, and J. -Y. Shih, "A study of Taiwan's issuer credit rating systems using support vector machines," *Expert Syst. Appl.*, vol. 30, no. 3, pp. 427-435, 2006.
[11] Y. -C. Lee, "Application of support vector machines to corporate credit rating prediction," *Expert Syst. Appl.*, vol. 33, no. 1, pp. 67-74, 2007.
[12] Y. S. Kwon, I. Han, and K. C. Lee, "Ordinal pairwise partitioning (OPP) approach to neural networks training in bond rating," *Intell. Syst. Account. Finance Manag.*, vol. 6, no. 1, pp. 23-40, 1997.

[13] H. Ahn, J. J. Ahn, H. W. Byun, and K. J. Oh, “A novel customer scoring model to encourage the use of mobile value added services,” *Expert Syst. Appl.*, vol. 38, no. 9, pp. 11693-11700, 2011.

[14] Y. -C. Wu, Y. -S. Lee, and J. -C. Yang, “Robust and efficient multiclass SVM models for phrase pattern recognition,” *Pattern Recogn.*, vol. 41, no. 9, pp. 2874-2889, 2008.

[15] A. C. Lorena, and A. C. P. L. F. de Carvalho, “Investigation of strategies for the generation of multiclass support vector machines,” in *New Challenges in Applied Intelligence Techniques*, N. T. Nguyen, and R. Katarzyniak, Eds. Berlin, Germany: Springer-Verlag, 2008, pp. 319-328.

[16] J. Friedman, “Another approach to polychtomous classification,” Technical Report, Stanford University, 1996.

[17] U. Kreβel, “Pairwise classification and support vector machines,” in *Advances in Kernal Methods: Support Vector Learning*, ch. 15, B. Schölkopf, C. Burges, and A. J. Smola, Eds. Cambridge, MA: MIT Press, 1999, pp. 255-268.

[18] J. Weston, and C. Watkins, “Support vector machines for multiclass pattern recognition,” in *Proc. 7th European Symp. Artificial Neural Networks*, Bruges, Belgium, 1999, pp. 219-224.

[19] C. -W. Hsu, and C. -J. Lin, “A simple decomposition method for support vector machines,” *Mach. Learn.*, vol. 46, nos. 1-3, pp. 291-314, 2002.

[20] F. E. H. Tay, and L. J. Cao, “Application of support vector machines in financial time series forecasting,” *Omega*, vol. 29, no. 4, pp. 309-317, 2001.

[21] C. -C. Chang, and C. -J. Lin, LIBSVM : a library for support vector machines, 2001.
Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm/