COMBINING B&B-BASED HYBRID FEATURE SELECTION AND THE IMBALANCE-ORIENTED MULTIPLE-CLASSIFIER ENSEMBLE FOR IMBALANCED CREDIT RISK ASSESSMENT

Jie SUN[a, b], Young-Chan LEE[b], Hui LI[a], Qing-Hua HUANG[a]

[a]*School of Economics and Management, Zhejiang Normal University, Jinhua, 321004 Zhejiang, China*
[b]*College of Management and Tourism, Dongguk University,
GyeongJu Campus, Gyeongju, 780-714 Gyeongbuk, Korea*

Abstract. An ideal model for credit risk assessment is supposed to select important features and process imbalanced data sets in an effective manner. This paper proposes an integrated method that combines B&B (branch and bound)-based hybrid feature selection (BBHFS) with the imbalance-oriented multiple-classifier ensemble (IOMCE) for imbalanced credit risk assessment and uses the support vector machine (SVM) and the multiple discriminant analysis (MDA) as the base predictor. BBHFS is a hybrid feature selection method that integrates the t-test and B&B with the k-fold cross-validation method to search for a satisfactory feature subset. The IOMCE divides majority samples into several subsets and then combines them with minority samples to construct several training sets for constructing a multiple-classifier ensemble model. We conduct main experiments using a 1:3 imbalanced corporate credit risk data set with continuous features and extended experiments using a 1:5 imbalanced data set with continuous features and a 1:3 imbalanced data set with discrete and nominal features. We combine no feature selection and five feature selection methods (the pure B&B, the factor analysis, the pure t-test, t-test & correlation analysis, and BBHFS) with single-classifier and the IOMCE to construct SVM and MDA models for an empirical comparison. When all features are continuous, the BBHFS-IOMCE method generally outperforms all the other methods. More specifically, BBHFS provides more stable and satisfactory results than the other feature selection methods, and compared with single-classifier models, IOMCE models can significantly enhance the recognition rate for minority samples while incurring a small reduction in the recognition rate for majority samples and maintaining an acceptable overall accuracy. When the features are almost discrete or nominal, the IOMCE method retains its ability to deal with an imbalanced data set, although the five feature selection methods have no significant advantages over no feature selection. This suggests that BBHFS is effective in retaining useful information when reducing the dimensionality of continuous features and that the BBHFS-IOMCE method is an important tool for imbalanced credit risk assessment.

Keywords: imbalanced credit risk assessment, imbalanced data set, hybrid feature selection, imbalance-oriented multiple-classifier ensemble.

JEL Classification: C38, G33, C45, G32.

Corresponding author Hui Li
E-mail: *lihuihit@gmail.com*

Routledge
Taylor & Francis Group

Introduction

To prevent new financial crises derived from credit crises, commercial banks as well as domain researchers place great emphasis on risk management, particularly credit risk management. In this regard, the assessment of credit risk is a particularly important issue. Research on credit risk assessment has received considerable attention from both scholars and practitioners, who typically employ a wide range of traditional statistical methods as well as new data-mining and artificial intelligence methods for credit risk modeling (Lahsasna *et al.* 2010) because the use of various techniques is generally expected to produce better modeling performance.

This study addresses three issues that have received little attention from previous research on credit risk assessment: the hybrid optimization of feature selection, the optimization of the feature extraction rate, and imbalanced modeling. The feature selection process entails the identification of an optimal feature subset based on a given data set with multiple features. Feature selection is used to increase model performance, reduce dimensionality, accelerate the speed of modeling, and reduce the computing time and cost. An increase in the size of a bank credit data set is likely to increase the amount of irrelevant information in the data set, and therefore feature selection can play an important role in credit decision modeling. However, previous research on credit risk modeling has focused mainly on the innovation of modeling algorithms (e.g. Zhou *et al.* 2010a; Wang *et al.* 2011) or the application of straightforward feature selection methods and thus paid little attention to performance improvements in modeling through feature selection. In addition, performances of some feature selection methods tend to vary with different feature extraction rates. In this regard, an important issue in the feature selection process is the determination of a suitable feature extraction rate for better classifier performance. More importantly, a credit data set is typically imbalanced, which implies that the number of samples with good credit often exceeds that of samples with bad credit. In this case, the results for the overall accuracy of a credit risk model trained and tested using imbalanced data sets may overestimate its performance. A high recognition rate for samples from the majority group may increase the model's overall recognition rate, but the recognition rate may remain low for those samples from the minority group. Therefore, it is necessary to address imbalanced data set to improve the ability of the credit decision model to identify the minority samples more accurately without incurring a significant decrease in its ability to identify the majority samples.

Based on the above discussion, this study proposes the BBHFS-IOMCE credit risk modeling approach by combining the B&B (branch and bound)-based hybrid feature selection (BBHFS) method with the imbalance-oriented multiple-classifier ensemble (IOMCE). The study employs the multiple-classifier ensemble to deal with imbalanced data sets. In addition, the study uses the support vector machine (SVM) and the multiple discriminant analysis (MDA) as the base predictor because the SVM is verified to be one of the most efficient models for assessing credit risk (Yu *et al.* 2010, 2011; Tseng *et al.* 2011; Bellotti *et al.* 2011; Xie *et al.* 2011; Xu *et al.* 2011; Liu *et al.* 2011; Wu *et al.* 2010; Lahsasna *et al.* 2010) and the MDA is a classical credit risk assessment method (Eisenbeis 1997). By combining the t-test and B&B, this study proposes a hybrid feature selection method for determining an optimal feature extraction rate driven by data, not by some subjective determination. The study provides

a comparison of credit risk assessment performance between BBHFS and the other three feature selection methods as well as no feature selection and between the IOMCE and the single classifier through experiments with three data sets.

The rest of this paper is organized as follows: Section 1 outlines the theoretical background. Section 2 presents the BBHFS-IOMCE methodology, and Section 3 describes the empirical experiments for the assessment of corporate credit risk. Section 4 provides extended experiments using two additional data sets.

1. Theoretical background

1.1. Previous research

This study classifies existing credit risk assessment models into three categories: credit risk assessment based on expert scoring that based on capital market theory or information economics, and that through multivariate classification analysis. Credit risk assessment based on expert scoring focuses mainly on evaluation elements and methods. Because this type of assessment embodies the idea of qualitative decision making, it facilitates subjective credit decisions, and decision outcomes are determined mainly by experts' experience and knowledge. Previous studies (e.g. Stiglitz, Wesiss 1981; Bester, Sreening 1985; Williamson 1986; Besanko, Thakor 1987; Hillier, Lbrahimo 1993) have extended the literature on credit risk assessment to include that based on capital market theory or information economics. This type of modeling is based on rigorous mathematical theories and the principles of information economics and reflects a better theoretical foundation. However, it cannot directly generate credit risk assessment results. Credit risk assessment through a multivariate classification analysis is the most effective and widely used method in the field of credit management. Eisenbeis (1997) employs the MDA for credit risk modeling, and Laitinen (1999) and Boguslauskas *et al.* (2011) verify the effectiveness of the logistic regression analysis in credit risk modeling. Finlay (2008) takes a utility approach to predict continuous financial measures such as contributions to profits in the context of the credit-scoring problem. Bellotti and Crook (2009) investigate the use of the survival analysis with macroeconomic variables for credit scoring. With the development of information sciences and decision sciences, previous studies have proposed some new techniques that overcome the limitations of traditional statistical models, including the fuzzy *k*-nearest neighbor algorithm (Laha 2007), neural networks (Khashman 2010; Angelini *et al.* 2008; Tsai, Wu 2008), the SVM (Chen *et al.* 2007; Huang 2009), rule extraction techniques based on neural networks (Baesens *et al.* 2003; Martens *et al.* 2007), ant colony optimization (Martens *et al.* 2010), and ensemble approaches (West *et al.* 2005; Paleologo *et al.* 2010), among others.

However, although the real-world problem of credit risk assessment is based on imbalanced data sets, the above studies have focused mainly on balanced data sets. In an imbalanced data set, samples with good credit generally exceed those with bad credit. Kennedy *et al.* (2009) compare the performance of a variety of one-class and two-class classifiers for imbalanced credit scoring, and find that the performance of the two-class classifiers falls off as the rate of class imbalance increases and one-class classifiers outperform two-class classifiers when the

data sets are extremely imbalanced with less than 2% default samples. Brown and Mues (2012) make an experimental comparison of classification algorithms such as logistic regression, MDA, neural networks, decision tree, gradient boosting, least square SVM, and random forest for imbalanced credit scoring data sets. They find that the random forest and gradient boosting classifiers perform better in an imbalanced credit scoring context, followed by the logistic regression and linear MDA. Resampling is one of the most widely used methods for dealing with class imbalance in the two-class classification problem. Garcia *et al.* (2012) empirically show that over-sampling the minority class consistently outperforms under-sampling the majority class when data sets are strongly imbalanced, whereas there are not significant differences for databases with a low imbalance. These studies provide some pioneering evidence for the problem of imbalance-oriented credit risk assessment modeling. However, it is still necessary to make a further study on some new and effective imbalance-oriented modeling methods for credit risk assessment.

In general, objects in data mining have many features, and some noisy and redundant features often occur, increasing the computing cost and the likelihood of misleading results. Guyon, Elisseeff (2003) provide an in-depth analysis on various variables and suggest that the feature selection method has many advantages such as facilitating the visualization and understanding of data, reducing measurement and storage requirements, reducing the training time, and improving prediction performance. However, previous studies of credit risk modeling have generally emphasized the innovation of modeling algorithms. For example, Yu *et al.* (2008, 2009, 2010) and Zhou *et al.* (2010a), employ a multistage neural network ensemble learning approach, the fuzzy group decision making model, the least squares SVM ensemble model and SVM-based multi-agent ensemble learning, respectively, for credit risk assessment. Zhou *et al.* (2010b) and Dong *et al.* (2010) propose the nearest-subspace method and a logistic regression approach with random coefficients, respectively. However, only a few studies have paid attention to the feature selection process in credit risk modeling. For example, Liu *et al.* (2005) and Wu *et al.* (2009) apply the SVM and the fuzzy integral SVM ensemble, respectively, for credit risk modeling, and include the feature selection method of factor analysis in the modeling process. In addition, Tsai (2009) finds that different feature selection methods result in different feature extraction rates and the corresponding classification performances. In this regard, optimizing the feature selection process and the determination of a suitable feature extraction rate is an important issue in credit risk assessment.

1.2. Feature selection methods

This study focuses on three feature selection methods: B&B, the t-test, and the factor analysis. The study considers B&B because it represents an optimization search in feature selection methods and is widely used as a global optimization method (Lawler, Wood 1966). Chen (2003) proposes an improved B&B algorithm for feature selection that can search for the same optimal solution faster than traditional ones. Tsai (2009) compares several feature selection methods for bankruptcy prediction and finds that the t-test and the factor analysis show good performance.

The basic idea of B&B is to divide a given problem (e.g. candidate features for credit risk assessment) into several subproblems (e.g. potential feature subsets). Then the algorithm

divides each subproblem continuously until it can no longer be decomposed or produce an optimal solution. B&B-based feature selection methods first construct a search tree of candidate features and then start searching from upper nodes to bottom ones through back-tracking. They search the tree from top to bottom and from right to left until they identify an optimal feature subset. This study uses the sum of the estimated Mahalanobis distance as an optimization criterion.

British psychologist Spearman (1904) proposes the factor analysis, a statistical technique for extracting common factors from multiple variables. It extracts fewer factors to describe relationships between multiple variables and classifies variables into various groups according to the dactyl of correlations between them. As a result, correlations between variables in the same group are strong and those between variables in different groups are weak. Here each group generates one factor, and therefore this analysis extracts fewer factors to reflect most information of the original data set. Factor extraction methods include the principal component analysis (PCA), the unweighted least squares method, the maximum likelihood method, and the image factor extraction method, among others. This study uses the PCA to extract important factors and determines the number of main factors and the corresponding eigenvector matrix based on the principle that the eigenvalue should exceed 1.

Based on test objects, the t-test can be classified into various types, including the single-sample t-test, the independent-sample t-test, and the paired-sample t-test. The single-sample t-test is used to test whether there is a significant difference between the mean of a single variable and a given constant. The independent-sample t-test examines the significance of mean differences between two independent samples and analyzes whether these two samples come from two populations with equal means. The paired-sample t-test is used to analyze the significance of mean differences between two paired samples. This study uses the independent-sample t-test to identify any significant mean differences between two groups for each feature and exclude uncorrelated features to reduce dimensionality.

2. BBHFS-IOMCE methodology

We propose a credit risk assessment model based on BBHFS-IOMCE by comprehensively considering the following: 1) the importance of optimizing feature selection in credit risk modeling; 2) the importance of determining an appropriate feature extraction rate; and 3) the reality that a credit data set typically consists of imbalanced samples. The proposed model uses a multiple-classifier ensemble to treat imbalanced samples and employs hybrid feature selection based on the t-test and B&B to determine an appropriate feature extraction rate and retain useful information.

2.1. BBHFS algorithm

For a data set with N features, suppose that the objective of feature selection is choosing D features from the initial N features to compose an optimal feature subset. Heijden *et al.* (2004) provide a detailed description of B&B, which can search for a globally optimal feature subset. However, the value of D is often determined by humans subjectively, and classification

models based on different D values tend to perform differently. In this regard, it is important to determine an appropriate D value or a suitable feature extraction rate (D/N). BBHFS first sets the upper and lower bounds of D according to t-test results, and by using a tenfold cross-validation method to divide the original training data set into 10 parts, it generates 10 groups of training and validation sets. Then, in each group of training and validation sets, it searches the feature subset that maximizes classifier performance. Finally, it calculates the frequency of each feature in the 10 feature subsets selected and reserves those features with frequency exceeding three to obtain a feature subset with an optimal D value. Figure 1 shows the BBHFS algorithm.
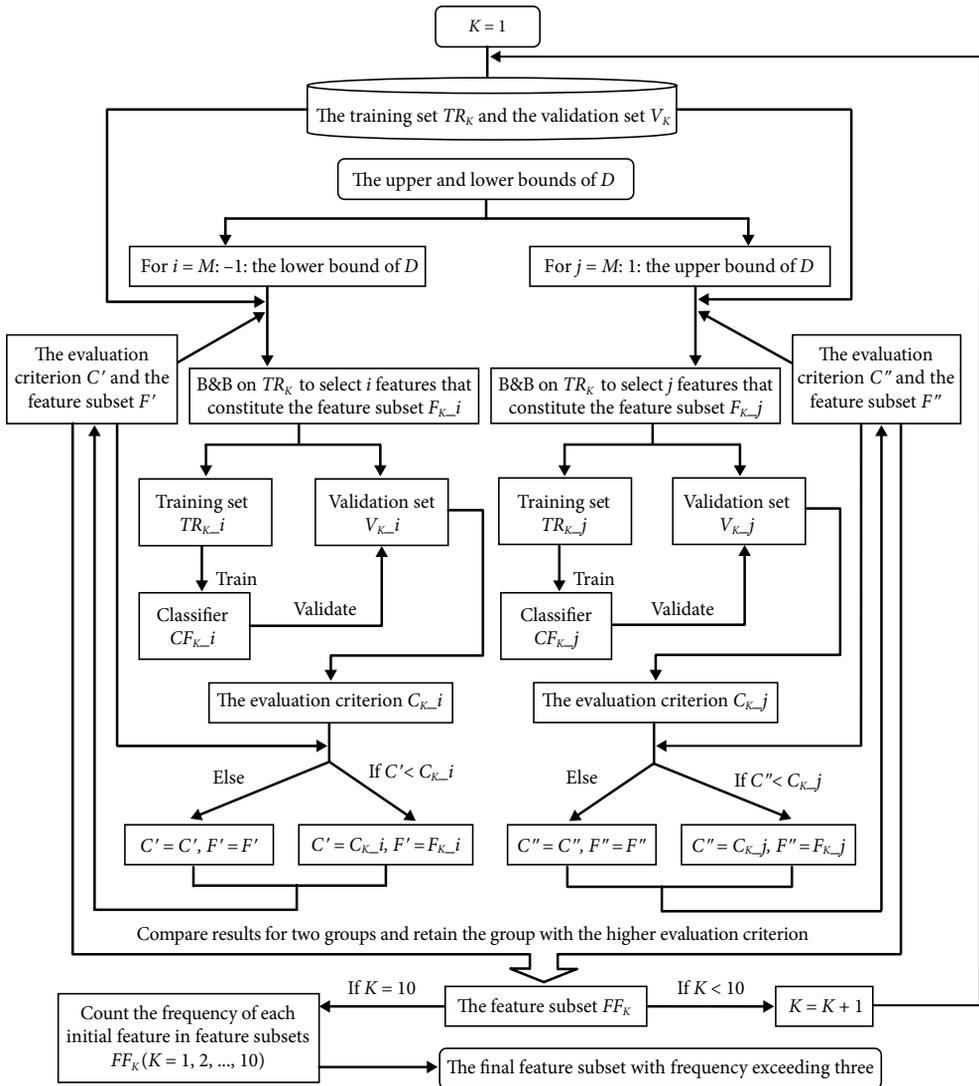

Fig. 1. BBHFS search process for an optimal feature subset

Algorithm:

(1) Select features that have significant differences between the two classes in the t-test to represent the initial training data set, denoted as *DS*.

(2) Set the upper bound of *D* as the number of features selected by t-test.

(3) Set the lower bound of *D* as 1.

(4) Determine the intermediate value of upper and lower bounds of *D* and denote it as *M*.

(5) Divide the initial training data set *DS* evenly into 10 subsets, which are denoted as $DS_1, DS_2, ..., DS_{10}$.

(6) For $K = 1, 2, ..., 10$

    1) Construct the training set $TR_K = DS_K$, and the validation set $V_K = DS - DS_K$.

    2) Search for the local optimal feature subset that makes the classifier trained on TR_K achieve the highest validation accuracy on V_K, by varying the value of *D* from *M* respectively to both the upper and lower bounds of *D*. Namely, implement the following processes ① and ② simultaneously:

    ① Search from the left side of *M: M* → the lower bound of *D*

        Initial setting: The left side's local optimal evaluation criterion *C′*= [];

                      The left side's local optimal feature subset *F′* = [].

        For $i = M, (M-1),...,$ the lower bound of *D*

        a) Use B&B on TR_K to select *i* features that constitute the feature subset $F_{K\_}i$;

        b) Obtain a training data subset $TR_{K\_}i$ from TR_K according to $F_{K\_}i$;

        c) Train the classifier $CF_{K\_}i$ on $TR_{K\_}i$;

        d) Obtain a validation data subset $V_{K\_}i$ from V_K according to $F_{K\_}i$;

        e) Verify $CF_{K\_}i$ on $V_{K\_}i$ to obtain the evaluation criterion $C_{K\_}i$;

        f) Update *C′* and *F′* according to the following rule:

            If $C′ < C_{K\_}i$, then $C′ = C_{K\_}i$ and $F′ = F_{K\_}i$;

    ② Search from the right side of *M: M* → the upper bound of *D*

        Initial setting: The right side's local optimal evaluation criterion *C″* = [];

                      The right side's local optimal feature subset *F″* = [];

        For $j = M, (M+1),...,$ the upper bound of *D*

        a) Use B&B on TR_K to select *j* features that constitute the feature subset $F_{K\_}j$;

        b) Obtain a training data subset $TR_{K\_}j$ from TR_K according to $F_{K\_}j$;

        c) Train the classifier $CF_{K\_}j$ on $TR_{K\_}j$;

        d) Obtain a validation data subset $V_{K\_}j$ from V_K according to $F_{K\_}j$;

        e) Verify $CF_{K\_}j$ on $V_{K\_}j$ to obtain the evaluation criterion $C_{K\_}j$;

        f) Update *C″* and *F″* according to the following rule:

            If $C″ < C_{K\_}j$, then $C″ = C_{K\_}j$ and $F″ = F_{K\_}j$;

    ③ Obtain the local optimal feature subset FF_k for the *k*-th iteration according to the following rule:

        If $C′ > C″$, then $FF_k = F′$, else $FF_k = F″$;

(7) After 10 iterations, collect 10 groups of feature subsets, namely $FF_1, FF_2, ..., FF_{10}$. Calculate the frequency of each feature in these feature subsets and reserve those features with frequency exceeding three. As a result, we obtain a final optimal feature subset whose corresponding *D* value is optimal.

From the above algorithm, we can find that BBHFS has the following advantages. After excluding features with no significant differences between two groups based on the t-test, BBHFS selects those features with higher frequency to compose an optimal feature subset by searching 10 groups of feature subsets that maximize classifier performance. In this way, it obtains an optimal D value that is not determined subjectively or arbitrarily. In addition, it uses a parallel searching strategy to search for an optimal D value. That is, it starts from the intermediate value of the upper and lower bounds of D and then searches from both sides simultaneously, reducing the computing time.

2.2. IOMCE

Data sets used for credit risk modeling are typically not in equilibrium, which means that samples of some categories account for a large portion of a data set, whereas others, a small portion. If such an imbalanced data set is used for single-classifier modeling, then the model's recognition rate for the majority group is likely to be higher than that for the minority group. However, more attention should be focused on identifying the minority group in some real-world application of credit risk assessment. Therefore, it is necessary to appropriately resolve the problem of imbalanced data. There are two major strategies for addressing this problem. A data-level strategy changes the distribution of data to reduce imbalance degree, and an algorithm-level strategy modifies algorithms so that they produce good modeling results even for imbalanced data (Chawla *et al.* 2004). From the data-level perspective, useful processes generally include oversampling, undersampling, and partitioning training set. Oversampling improves classification performance for the minority group by adding minority samples (Maloof 2003). For instance, Chawla *et al.* (2002) propose the synthetic minority over-sampling technique. Undersampling deletes some majority samples to improve classification performance for the minority group. For example, Kubat and Matwin (1997) propose a one-sided selection algorithm. Oversampling and undersampling both have some limitations. Oversampling tends to increase minority samples by copying samples without adding any new information on the minority group and thus may lead to overfitting. Undersampling may cause the loss of some important information because it reduces some majority samples.

The training set partition method addresses imbalanced data by dividing the training data set. It divides majority samples from the training data set into a series of disjointed subsets based on the ratio of majority samples to minority ones and then combines the latter with each of the above subsets separately to generate various training subsets. A combined classification method using multiple classifiers trained independently from various training subsets can outperform oversampling or undersampling methods (Yan *et al.* 2003). From the algorithm-level perspective, useful strategies include the cost-sensitive learning method and the classifier ensemble method. The cost-sensitive learning method sets misclassification costs that vary with different groups. That is, it sets higher costs for the minority group than for the majority group. In this way, it enhances the recognition rate for the minority group. However, it is difficult to determine misclassification costs for minority and majority samples in credit risk assessment. The classifier ensemble method combines a variety of base classifiers through a certain ensemble rule to generate combined classification results,

enhancing overall classification performance for an imbalanced data set. Although there are other methods, including active learning (Constantinopoulos, Likas 2008), random forests (Chen *et al.* 2004), and the subspace method (Ahn *et al.* 2007), the training set partition method and the classifier ensemble method can better address imbalanced data sets (Yei *et al.* 2009). Therefore, the present paper uses the IOMCE method, which combines these two methods to address imbalanced data sets for credit risk assessment. Figure 2 shows the IOMCE process for binary classification.
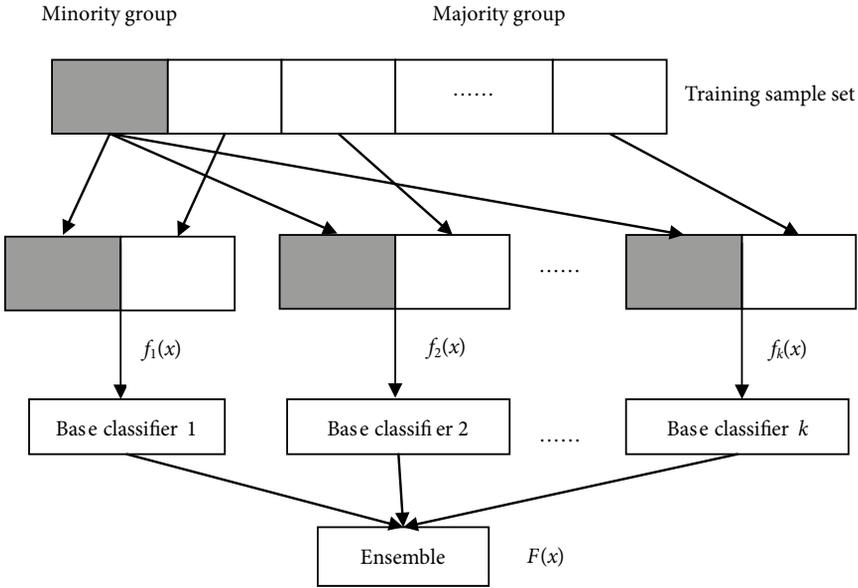

Fig. 2. IOMCE process

Let the ratio of majority samples to minority ones be θ. The partition ratio of majority samples k is set according to the value of θ. In detail, if θ is odd, then the partition ratio of majority samples is equal to θ, that is, $k = θ$, and simple majority voting can be used as the ensemble rule. On the other hand, if θ is even and $k = θ$, then simple majority voting may result in the same votes for the two groups, causing a decision dilemma. To avoid such a situation, we can employ two strategies. First, we can let $k = θ – 1$. In this case, the number of base classifiers in the ensemble is still odd, but the number of minority samples and that of majority samples in the training data set are still imbalanced to some extent, although this imbalance is much less pronounced (i.e. close to a balanced data set). Second, we can use weighted majority voting instead of simple majority voting as the ensemble rule. In this case, training accuracy can serve as the weight for the vote from each base classifier, which can make the weighted votes for the two groups different in most cases. Even when the same weighted votes are obtained for the two groups, which happens with little possibility, the target sample can be recognized as having bad credit for the purpose of controlling credit risk.

3. Empirical experiments for the assessment of corporate credit risk

3.1. Experimental design

For feature selection, we consider no feature selection as well as the following four feature selection methods: pure B&B, the factor analysis, the pure t-test, and BBHFS. We randomly divide each group of samples into two parts so that two thirds of samples are used for training and the rest, for testing. Because any difference in the partition of a data set can lead to different experimental results, we repeat the whole experiment for totally 100 times to avoid a biased analysis. That is, we partition the whole data set into two subsets (one for training and the other for testing) for totally 100 times. We construct a credit risk assessment model for each training data set and test it by using the corresponding testing data set.

Because training and testing data sets are imbalanced, we use the IOMCE method for credit risk modeling. In the experimental data set, the ratio of highly risky firms to normal ones is 1:3. Therefore, for the training set, we trisect normal firms into three parts and combine each of them with highly risky firms to generate three training subsets. We use them to train three base classifiers and then integrate their outcomes by majority voting.

We conduct empirical experiments by using Matlab with LibSVM (Fan *et al.* 2005) and PRTool (Heijden *et al.* 2004) and use SPSS to replace missing data and to conduct the factor analysis and the t-test.

3.2. Sample data and feature selection

We collect publicly traded firms that are highly risky in the years 2001–2010 and show the sample with financial ratios for year t-2. Here t denotes the year in which a publicly traded firm receives special treatment (ST) from the Shanghai Stock Exchange or the Shenzhen Stock Exchange. We define those publicly traded firms in China that receive ST for a bad financial condition in year t and once increased their long-term bank loan in year t-2 as highly risky samples. We match highly risky samples with low risky ones. We select three matched samples for each highly risky sample based on the following criteria: 1) they belong to the same (or similar) industry as the highly risky sample; 2) they have no ST experience; and 3) they show good operating performance. The matching ratio is 1:3. Finally, we obtain 107 highly risky firms and 321 normal ones for credit risk assessment (a total of 428 sample firms). All these firms are medium-sized or large firms because their registered capital is at least CNY 50 million. We use the credit risk assessment model to predict these two types of firms. If the model can successfully predict highly risky firms, then banks can adjust their credit decisions and reduce their loss from credit risk.

The initial feature set consists of 16 financial ratios. All these ratios belong to continuous features and indicate a firm's solvency, profitability, operating capability, or development capacity (Table 1). We collect values of the 16 financial ratios for year t-2 for each sample and obtain the original experimental data set. Missing values require careful handling in credit risk assessment (Florez-Lopez 2010). Missing values for each feature are replaced with the average value of all samples in the same group. The factor analysis extracts six factors from the original 16 features. The features selected by other methods are marked in Table 1. The results of independent-

Table 1. Candidate features of financial ratios and feature selection results

Type	Variable	Feature name	t-test	t-test & CA	B&B5	BBHFS
Solvency	V1	Current ratio	√			
	V2	Cash to current debt ratio	√	√	√	√
	V3	Debt to asset ratio				
	V4	Debt to equity ratio	√	√		
	V5	Debt to tangible assets ratio				
Profitability	V6	Return on equity	√			
	V7	Return on total assets	√	√	√	√
	V8	Operating profit ratio			√	
	V9	Profit to cost ratio	√	√	√	√
Operating capability	V10	Accounts receivable turnover				
	V11	Inventory turnover				
	V12	Current assets turnover	√			
	V13	Total assets turnover	√	√		√
Development capacity	V14	Operating income growth rate	√	√		
	V15	Net profit growth rate	√	√	√	√
	V16	Total asset expansion rate	√	√		

sample t-test show that 11 features have significant differences between the high- and low-risk groups, and only five features are excluded. With the number of selected features set as 5, the five features selected by pure B&B (B&B5) include cash to current debt ratio, return on total assets, operating profit ratio, profit to cost ratio, and net profit growth rate. BBHFS automatically selects the following 5 features: cash to current debt ratio, return on total assets, profit to cost ratio, total assets turnover, and net profit growth rate, which all show significant differences in the t-test. Furthermore, we utilize the Pearson correlation analysis (CA) to validate the correlation of each pair of financial ratios selected by the pure t-test, B&B5, and BBHFS. As a result, only three pairs of financial ratios selected by the pure t-test show the correlation coefficients higher than 0.6. Namely, that of current ratio (V1) and debt to equity ratio (V4) is 0.65, that of return on equity (V6) and return on total assets (V7) is 0.61, and that of current assets turnover (V12) and total assets turnover (V13) is 0.72. However, none pair of financial ratios selected by B&B5 and BBHFS has a correlation coefficient over 0.6. Therefore, the pure t-test tends to retain too many features with redundant information, and we can combine it with CA to form the feature selection method of t-test & CA to further eliminate redundant ratios. Although the pure B&B does not select features that show high correlation between each other, it has difficulty in determining the number of selected features. Simultaneously, 3 out of 5 features selected by the B&B5 method are profitability ones and they do not cover operating capability at all, which may affect the performance of credit risk assessment. By contrast, BBHFS can not only select a feature subset with an optimal feature number, but also remove the redundant and relatively less important features, which have high correlation with some other important features. The features selected by BBHFS cover the whole dimensions of a company's solvency, profitability, operating capability, and development capacity, which should be emphasized by firms for healthy development and good credit rating.

3.3. Basic model

We apply the SVM and the MDA to train classifiers because they respectively belong to the new machine learning classification algorithm and the classical statistical classification method, respectively. The SVM, proposed by Vapnik (1996), is a machine learning algorithm based on statistical learning theory. The SVM, which is based on statistical VC-dimension theory and the structural risk minimization principle, outperforms other artificial intelligence methods in terms of modeling and generalization and can better deal with small samples. The SVM is a very useful tool for credit risk assessment (Crook *et al.* 2007; Bellotti, Crook 2008; Cimpoeru 2011). Previous studies have indicated that the SVM algorithm with the radial basis function (RBF) has better generalization ability (Huang *et al.* 2004), and therefore we employ this type of SVM algorithm. The RBF SVM has two parameters, namely the tuning parameter C and the kernel parameter γ, and the choice of parameter values is critical to its classification performance. We combine the fivefold cross-validation method with the grid search technique to select optimal parameter values. The optimal values for C and γ are searched in the ranges $[2^{-8}, 2^{-7}, 2^{-6}, \ldots, 2^9, 2^{10}]$ and $[2^{-10}, 2^{-9}, 2^{-8}, \ldots, 2^7, 2^8]$, respectively. The MDA, proposed by Fisher (1936), is widely known as a fast and effective classification method. No parameters need to be set in advance for MDA modeling, and it can produce stable and acceptable performance in credit risk assessment (Chijoriga 2011; Hui, Sun 2006).

3.4. Evaluation measures

Accuracy, sensitivity and specificity are used as evaluation measures of classification performances. To define these evaluation measures clearly, we use the confusion matrix as a tool for analyzing the recognition ability of a classifier. The confusion matrix for binary classification is illustrated in Table 2.

Table 2. Confusion matrix for binary classification

	Positive prediction	Negative prediction
Positive group	TP	FN
Negative group	FP	TN

Here the positive group includes highly risky minority samples, and the negative one, low risky majority samples. In Table 2, TP (true positive) denotes the number of samples predicted as the positive when they are truly positive; FN (false negative) denotes the number of samples predicted as the negative when they are truly positive; FP (false positive) denotes the number of samples predicted as the positive when they are truly negative; and TN (true negative) denotes the number of samples predicted as the negative when they are truly negative. The performance evaluation measures are defined in Table 3.

Accuracy measures a classifier's overall classification capacity, sensitivity measures the capacity of a classifier to identify positive samples, and specificity measures the capacity of a classifier to identify negative samples.

Table 3. Performance evaluation measures

Measure	Calculation method
Accuracy	*Accuracy = (TP+TN) / (TP+FP+TN+FN)*
Sensitivity	*Sensitivity = TP / (TP+FN)*
Specificity	*Specificity = TN / (TN+FP)*

3.5. Empirical results and analysis

3.5.1. Results of evaluation measures for the corporate credit risk data set

Table 4 shows the means of accuracy, sensitivity, and specificity for single-classifier and IOMCE models using different feature selection methods for the corporate credit risk data set. B&B5 and B&B10 indicate the use of pure B&B for selecting 5 and 10 features, respectively, for credit risk modeling.

Table 4. Means of accuracy, sensitivity, and specificity for single-classifier and IOMCE models using different feature selection methods for the corporate credit risk data set (%)

Measure	Method	No selection	B&B5	B&B10	Factor analysis	t-test	t-test & CA	BBHFS
Accuracy	Single SVM	<u>83.13</u>	86.59	85.61	84.99	85.98	86.25	**87.22**
	SVM ensemble	<u>76.77</u>	86.66	84.26	80.17	85.50	85.96	**86.87**
	Single MDA	84.03	**87.36**	84.35	<u>81.57</u>	84.56	85.71	86.04
	MDA ensemble	81.16	**86.58**	82.20	<u>78.40</u>	81.46	83.29	84.95
Sensitivity	Single SVM	<u>45.89</u>	64.39	58.69	63.50	61.83	62.28	65.75
	SVM ensemble	<u>70.56</u>	**82.28**	80.58	80.14	80.61	80.53	81.00
	Single MDA	73.92	73.97	**75.61**	74.86	73.14	74.53	<u>72.64</u>
	MDA ensemble	80.08	77.92	**80.17**	<u>77.42</u>	78.69	79.86	77.86
Specificity	Single SVM	95.66	94.06	94.66	<u>92.21</u>	94.10	94.32	94.45
	SVM ensemble	<u>78.86</u>	88.14	85.50	80.19	87.14	87.79	**88.84**
	Single MDA	87.43	**91.87**	87.29	<u>83.83</u>	88.40	89.48	90.55
	MDA ensemble	81.52	**89.50**	82.89	<u>78.73</u>	82.39	84.45	87.34

3.5.2. Comparison between BBHFS and the other feature selection methods

Table 5 shows the results of mean comparison between BBHFS and the other feature selection methods for the corporate credit risk data set. In the table, Y (N) indicates that BBHFS has a higher (lower) mean of certain evaluation measure than the other feature selection methods. We conduct the nonparametric Wilcoxon test for two related samples to determine whether there are significant differences between BBHFS and the other methods. The values for Z-statistic and corresponding significance levels are also listed in Table 5.

Table 5. Results of the Wilcoxon test for mean comparison between BBHFS and the other feature selection methods for the corporate credit risk data set

Measure	Method	BBHFS > No selection	BBHFS > B&B5	BBHFS > B&B10	BBHFS > Factor analysis	BBHFS > t-test	BBHFS > t-test & CA
Accuracy	Single SVM	**Y: −8.45*****	**Y: −3.18*****	**Y: −4.56*****	**Y: −6.67*****	**Y: −5.63*****	**Y: −3.10*****
	SVM ensemble	**Y: −8.69*****	**Y: −1.17*****	**Y: −5.73*****	**Y: −8.64*****	**Y: −4.82*****	**Y: −2.48****
	Single MDA	**Y: −6.63*****	<u>N: −6.79***</u>	**Y: −5.81*****	**Y: −8.41*****	**Y: −5.14*****	Y: −0.98
	MDA ensemble	**Y: −8.00*****	<u>N: −6.31***</u>	**Y: −7.13*****	**Y: −8.68*****	**Y: −7.64*****	**Y: −4.29*****
Sensitivity	Single SVM	**Y: −8.69*****	**Y: −2.04****	**Y: −5.08*****	**Y: −2.29****	**Y: −5.15*****	**Y: −2.86*****
	SVM ensemble	**Y: −8.07*****	<u>N: −4.05***</u>	Y: −0.34	Y: −1.18	Y: −0.65	Y: −0.32
	Single MDA	<u>N: −1.99**</u>	<u>N: −3.33***</u>	<u>N: −5.03***</u>	<u>N: −2.73***</u>	N: −0.91	N: −1.21
	MDA ensemble	<u>N: −3.25***</u>	N: −0.16	<u>N: −4.02***</u>	Y: −0.43	N: −1.46	N: −1.47
Specificity	Single SVM	<u>N: −3.72***</u>	**Y: −2.32****	N: −0.83	**Y: −6.40*****	Y: −1.62	Y: −0.58
	SVM ensemble	**Y: −8.31*****	**Y: −2.81*****	**Y: −5.33*****	**Y: −8.64*****	**Y: −4.46*****	**Y: −1.86***
	Single MDA	**Y: −7.44*****	<u>N: −6.74***</u>	**Y: −7.87*****	**Y: −8.51*****	**Y: −5.99*****	**Y: −2.27****
	MDA ensemble	**Y: −8.09*****	<u>N: −7.05***</u>	**Y: −7.84*****	**Y: −8.64*****	**Y: −7.95*****	**Y: −4.43*****

Note: *** and ** indicate significance at the levels of 1% and 5%, respectively.

By comparing the values in Table 4 among different feature selection methods, we can find that no feature selection method always outperforms the others for each evaluation measure and for each classification algorithm. Therefore, the purpose of feature selection study is finding a satisfactory method instead of an optimal one, for a specific domain. As shown in Table 5, in the 72 times of comparison with the other feature selection methods, BBHFS performs significantly better for 46 times and worse for 12 times, with no significant difference for the remaining 14 times. This suggests that BBHFS is a satisfactory feature

selection method for credit assessment modeling. For a more objective evaluation, more detailed comparison between BBHFS and each of the other feature selection methods are explained as follows.

(1) *BBHFS vs. no feature selection*: In the 12 times of comparison with no feature selection, BBHFS obtains significantly better performance for 9 times and significantly worse performance for 3 times. Therefore, BBHFS significantly outperforms no feature selection.

(2) *BBHFS vs. pure B&B*: In the 12 times of comparisons with B&B5, BBHFS obtains significantly better performance for 5 times and significantly worse performance for 6 times, with no significant difference for 1 time. In addition, in the 12 times of comparisons with B&B10, BBHFS obtains significantly better performance for 8 times and significantly worse performance for 2 times, with no significant difference for 2 time. This suggests that the performance of pure B&B varies widely according to the number of selected features and that properly setting the number of selected features for pure B&B can produce more satisfactory results. Because the optimal number of selected features is unknown beforehand, this can be a key bottleneck when applying the pure B&B feature selection method to real-world problems. By contrast, BBHFS can determine the number of selected features automatically in a data-driven manner. Although it cannot ensure optimal performance, it shows very satisfactory performance. Together, these results suggest the superiority of BBHFS over pure B&B.

(3) *BBHFS vs. the factor analysis*: In the 12 times of comparison with the factor analysis, BBHFS obtains significantly better performance for 9 times and significantly worse performance for 1 time, with no significant difference for 2 times. Therefore, BBHFS significantly outperforms the factor analysis.

(4) *BBHFS vs. the t-test*: In the 12 times of comparison with the t-test, BBHFS obtains significantly better performance for 8 times, with no significant difference for 4 times. Therefore, BBHFS significantly outperforms the t-test.

(5) *BBHFS vs. the t-test & CA*: In the 12 times of comparison with the t-test & CA, BBHFS obtains significantly better performance for 7 times, with no significant difference for 5 times. Therefore, BBHFS significantly outperforms the t-test & CA.

Stability comparison among different feature selection methods are shown in Table 6. We calculate the coefficient of variation for accuracy, sensitivity, and specificity of single-classifier and IOMCE models for the corporate credit risk data set and make a comparison of them between B&B-based hybrid method and the other methods. In Table 6, Y indicates that BBHFS has a lower or equal coefficient of variation compared with other feature selection methods, and N indicates that BBHFS has a higher coefficient of variation. The results show that the features selected by BBHFS have better stability of credit risk assessment than no feature selection, B&B5, B&B10, factor analysis, and pure t-test, and almost equal stability to t-test & CA. More specifically, BBHFS produces lower or equal coefficients of variation for at least 6 times in the total 12 times of comparison with each of the other methods. By contrast, the features selected by no feature selection, factor analysis, and pure t-test show relatively worse stability in credit risk assessment. These results suggest that BBHFS is also a desirable feature selection method in terms of performance stability.

Table 6. Coefficient of variation for accuracy, sensitivity, and specificity of single-classifier and IOMCE models using different feature selection methods for the corporate credit risk data set

Measure	Method	No selection	B&B5	B&B10	Factor analysis	t-test	t-test & CA	BBHFS
Accuracy	Single SVM	**Y:** 0.031	**Y:** 0.027	N: 0.024	**Y:** 0.032	**Y:** 0.029	N: 0.024	0.027
	SVM ensemble	**Y:** 0.053	**Y:** 0.035	N: 0.033	**Y:** 0.035	**Y:** 0.033	N: 0.027	0.030
	Single MDA	**Y:** 0.035	**Y:** 0.027	**Y:** 0.036	**Y:** 0.036	**Y:** 0.033	**Y:** 0.030	0.026
	MDA ensemble	**Y:** 0.040	N: 0.028	**Y:** 0.037	**Y:** 0.044	**Y:** 0.044	**Y:** 0.033	0.033
Sensitivity	Single SVM	**Y:** 0.188	**Y:** 0.129	**Y:** 0.150	N: 0.126	N: 0.112	N: 0.116	0.128
	SVM ensemble	**Y:** 0.117	**Y:** 0.071	**Y:** 0.085	**Y:** 0.088	**Y:** 0.080	**Y:** 0.089	0.071
	Single MDA	N: 0.103	**Y:** 0.119	N: 0.103	N: 0.096	N: 0.105	**Y:** 0.140	0.115
	MDA ensemble	N: 0.094	N: 0.106	N: 0.099	N: 0.092	N: 0.105	**Y:** 0.123	0.107
Specificity	Single SVM	**Y:** 0.039	**Y:** 0.030	**Y:** 0.029	**Y:** 0.034	**Y:** 0.031	N: 0.023	0.029
	SVM ensemble	**Y:** 0.085	**Y:** 0.044	**Y:** 0.045	**Y:** 0.048	**Y:** 0.046	N: 0.040	0.041
	Single MDA	**Y:** 0.043	N: 0.028	**Y:** 0.047	**Y:** 0.049	**Y:** 0.045	**Y:** 0.040	0.034
	MDA ensemble	N: 0.050	N: 0.042	N: 0.048	**Y:** 0.063	**Y:** 0.057	N: 0.051	0.052
Number of **Y**		9	8	7	9	9	6	–

3.5.3. Comparison between IOMCE and single-classifier models

Based on Table 4, we calculate the difference between specificity and sensitivity for the corporate credit risk data set and show the results in Table 7. For a clearer understanding of the relationships between single-classifier and IOMCE models in terms of their performance evaluation measures, we provide graphs in Figure 3 based on the results in Table 4.

Table 7. Difference between specificity and sensitivity for the corporate credit risk data set

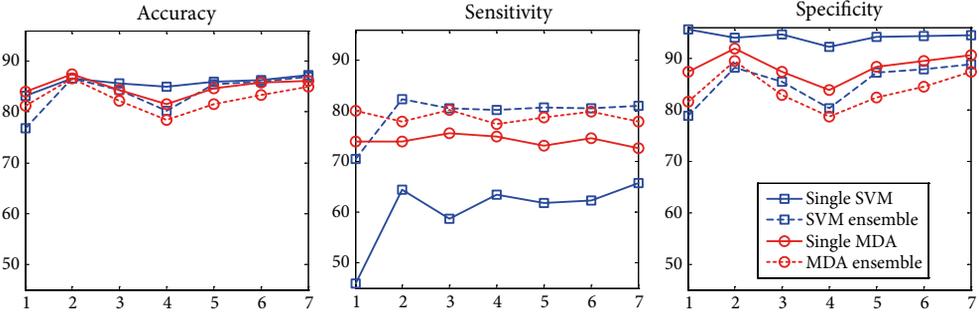| Method	No selection	B&B5	B&B10	Factor analysis	t-test	t-test & CA	BBHFS
Single SVM	49.77	29.67	35.97	28.71	32.27	32.04	28.7
SVM ensemble	8.3	5.86	4.92	0.05	6.53	7.26	7.84
Single MDA	13.51	17.9	11.68	8.97	15.26	14.95	17.91
MDA ensemble	1.44	11.58	2.72	1.31	3.70	4.59	9.48

Fig. 3. Graphs of performance evaluation measures of single-classifier and IOMCE models
for the corporate credit risk data set

As shown in Table 4 and Figure 3, single-SVM and single-MDA show relatively high accuracy and specificity but low sensitivity. In addition, Tables 4 and 7 show large difference between specificity and sensitivity for single-SVM and single-MDA models with different feature selection methods, which can be explained by the imbalanced experimental data set. More specifically, single-SVM and single-MDA models trained on an imbalanced data set tend to produce high recognition rates for majority samples, which results in high degree of specificity and drives up the overall accuracy. However, single-SVM and single-MDA models produce relatively low recognition rates for minority samples, corresponding to relatively low degree of sensitivity. As shown in Table 7 and Figure 3, this classification phenomenon for an imbalanced data set is more pronounced for single-SVM models than for single-MDA models. That is, single-SVM models show even lower degrees of sensitivity and even larger differences between specificity and sensitivity than those for single-MDA models.

With the employment of IOMCE to process imbalanced data set, the differences between specificity and sensitivity reduce significantly, as shown by Table 7. The differences between specificity and sensitivity are significantly lower for IOMCE than for single-classifier models, and the reductions are even more evident for SVM models than for MDA models.

As shown in Table 4, the employment of the IOMCE method does not increase overall accuracy for most SVM and MDA models with different feature selection methods, and in some cases, it even reduces it slightly. However, the employment of the IOMCE method sharply increases sensitivity with acceptable reductions in specificity for all models. Table 8 shows the differences in accuracy, sensitivity, and specificity between IOMCE and single-classifier models for the corporate credit risk data set. The results clearly indicate that processing an imbalanced data set with the IOMCE method substantially improves the classifier's ability to identify minority samples (the group with a high level of credit risk) while incurring a smaller decrease in its ability to identify majority samples (the group with a low level of credit risk) for all models except for MDA with the factor analysis and the t-test.

Table 8. Differences in accuracy, sensitivity, and specificity between IOMCE and single-classifier models for the corporate credit risk data set

Measure	Method	No selection	B&B5	B&B10	Factor analysis	t-test	t-test & CA	BBHFS
Accuracy	SVM	–6.36	0.07	–1.35	–4.82	–0.48	–0.29	–0.35
	MDA	–2.87	–0.78	–2.15	–3.17	–3.10	–2.42	–1.09
Sensitivity	SVM	24.67	17.89	21.89	16.64	18.78	18.25	15.25
	MDA	6.16	3.95	4.56	2.56	5.55	5.33	5.22
Specificity	SVM	–16.8	–5.92	–9.16	–12.02	–6.96	–6.53	–5.61
	MDA	–5.91	–2.37	–4.4	–5.1	–6.01	–5.03	–3.21

These results suggest that the IOMCE method should be used to address imbalanced data sets in credit risk modeling. In particular, combining BBHFS with the IOMCE can produce stable and satisfactory performance in credit risk modeling based on an imbalanced data set. Overall, the BBHFS-IOMCE method is more effective than all other methods, suggesting that it is a useful credit risk assessment tool for publicly traded firms in China.

4. Extended experiments for financial distress prediction and personal credit risk assessment

4.1. Data

We conduct extended experiments using two additional data sets to validate whether the proposed BBHFS-IOMCE approach is effective for other data sets. The first data set for financial distress prediction (FDP) is 1:5 imbalanced. It includes 384 samples, among which 64 are distressed samples and 320 are healthy ones. This data set is an imbalanced version processed from the original balanced data set in Sun and Li (2011) and has 42 initial features that are all continuous. This data set is similar to the above corporate credit risk data set in that all their features are continuous, and the difference is that the FDP data set has much more initial features. The second data set is derived from a UCI data set (German personal credit risk) provided by Hofmann and Stat (1994). The original data set consists of 300 samples with bad credit and 700 samples with good credit. By randomly duplicating 200 samples with good credit, we obtain a new 1:3 imbalanced data set with a total of 1,200 samples. This data set is different from the corporate credit risk data set and the FDP data set in that it includes many discrete and nominal features.

4.2. Experimental results and analysis for the FDP data set

4.2.1. Results of evaluation measures for the FDP data set

Table 9 shows the means of accuracy, sensitivity, and specificity for single-classifier and IOMCE models using different feature selection methods for the FDP data set. B&B10 and B&B20 indicate the use of pure B&B for selecting 10 and 20 features, respectively, for FDP modeling.

Table 9. Means of accuracy, sensitivity, and specificity for single-classifier and IOMCE models using different feature selection methods for the FDP data set (%)

Measure	Method	No selection	B&B10	B&B20	Factor analysis	t-test	t-test & CA	BBHFS
Accuracy	Single SVM	90.36	**91.10**	90.86	90.13	<u>89.24</u>	90.18	90.80
	SVM ensemble	85.74	**89.64**	88.09	85.49	<u>81.68</u>	87.35	88.68
	Single MDA	<u>88.18</u>	91.98	90.95	89.53	90.04	89.75	**92.37**
	MDA ensemble	<u>79.43</u>	89.22	85.63	84.98	83.33	84.29	**89.33**
Sensitivity	Single SVM	56.59	**59.18**	55.50	57.00	<u>47.55</u>	55.32	57.55
	SVM ensemble	78.68	**81.86**	79.82	<u>77.59</u>	79.55	78.55	80.64
	Single MDA	<u>62.59</u>	69.64	65.41	68.41	64.41	64.86	**70.45**
	MDA ensemble	<u>69.00</u>	77.50	76.55	76.36	76.45	77.14	**79.41**
Specificity	Single SVM	97.30	97.66	**98.13**	<u>96.94</u>	97.81	97.35	97.64
	SVM ensemble	87.20	**91.23**	89.79	87.11	<u>82.12</u>	89.16	90.34
	Single MDA	<u>93.44</u>	96.58	96.21	93.87	95.31	94.87	**96.88**
	MDA ensemble	<u>81.58</u>	**91.63**	87.50	86.75	84.75	85.77	91.37

4.2.2. Comparison between BBHFS and the other feature selection methods

Table 10 shows the results of mean comparison between BBHFS and the other feature selection methods for the FDP data set as well as those of the nonparametric Wilcoxon test. In the 72 times of comparison with the other feature selection methods, BBHFS performs significantly better for 54 times and worse for 4 times, with no significant difference for the remaining 14 times. More specifically, BBHFS shows significantly better performance than no feature selection, B&B10, B&B20, the factor analysis, the pure t-test, and t-test & CA for 11, 4, 8, 11, 10, and 10 times, respectively, and worse performance for 0, 3, 1, 0, 0, and 0 times. These results provide support for the conclusion drawn from the corporate credit risk data set that BBHFS is a satisfactory feature selection method for continuous features.

Table 11 shows the coefficient of variation for accuracy, sensitivity, and specificity for single-classifier and IOMCE models using different feature selection methods for the FDP data set, as well as the result of stability comparison between BBHFS and the other methods. BBHFS produces lower coefficients of variation for at least 7 times in the total 12 times of comparison with each of the other methods. These results provide support for the conclusion that BBHFS is more stable than the other feature selection methods.

Table 10. Results of the Wilcoxon test for mean comparison between BBHFS and the other feature selection methods for the FDP data set

Measure	Method	BBHFS > No selection	BBHFS > B&B10	BBHFS > B&B20	BBHFS > Factor analysis	BBHFS > t-test	BBHFS > t-test & CA
Accuracy	Single SVM	**Y: −2.28****	N: −2.26**	N: −0.28	**Y: −3.39*****	**Y: −6.08*****	**Y: −2.48****
	SVM ensemble	**Y: −7.47*****	N: −5.09***	**Y: −1.79***	**Y: −7.45*****	**Y: −8.67*****	**Y: −2.69*****
	Single MDA	**Y: −8.59*****	**Y: −3.43*****	**Y: −6.66*****	**Y: −8.61*****	**Y: −7.96*****	**Y: −7.50*****
	MDA ensemble	**Y: −8.69*****	Y: −0.69	**Y: −8.06*****	**Y: −8.28*****	**Y: −8.61*****	**Y: −7.96*****
Sensitivity	Single SVM	Y: −0.93	N: −1.64	**Y: −2.69*****	Y: −0.69	**Y: −7.05*****	**Y: −1.73***
	SVM ensemble	**Y: −2.95*****	N: −1.37	Y: −0.79	**Y: −3.49*****	Y: −1.51	**Y: −2.07****
	Single MDA	**Y: −7.54*****	**Y: −1.66***	**Y: −6.39*****	**Y: −3.44*****	**Y: −7.06*****	**Y: −3.76*****
	MDA ensemble	**Y: −7.31*****	**Y: −2.62*****	**Y: −3.26*****	**Y: −3.56*****	**Y: −3.49*****	Y: −1.29
Specificity	Single SVM	**Y: −1.79***	N: −0.04	N: −2.56**	**Y: −2.96*****	N: −1.07	Y: −0.79
	SVM ensemble	**Y: −6.62*****	N: −4.34***	Y: −1.37	**Y: −6.76*****	**Y: −8.51*****	**Y: −1.74***
	Single MDA	**Y: −8.39*****	**Y: −2.40****	**Y: −4.03*****	**Y: −8.23*****	**Y: −7.06*****	**Y: −6.27*****
	MDA ensemble	**Y: −8.69*****	N: −1.23	**Y: −7.97*****	**Y: −8.21*****	**Y: −8.61*****	**Y: −7.73*****

Note: ***, **, and * indicate significance at the levels of 1%, 5%, and 10%, respectively.

Table 11. Coefficient of variation for accuracy, sensitivity, and specificity for single-classifier and IOMCE models using different feature selection methods for the FDP data set

Measure	Method	No selection	B&B5	B&B10	Factor analysis	t-test	t-test & CA	BBHFS
Accuracy	Single SVM	**Y: 0.020**	N: 0.018	N: 0.019	N: 0.019	**Y: 0.021**	N: 0.018	0.020
	SVM ensemble	**Y: 0.041**	**Y: 0.029**	**Y: 0.032**	**Y: 0.035**	**Y: 0.047**	**Y: 0.036**	0.027
	Single MDA	**Y: 0.029**	**Y: 0.020**	**Y: 0.022**	**Y: 0.022**	**Y: 0.026**	**Y: 0.019**	0.019
	MDA ensemble	**Y: 0.042**	**Y: 0.027**	**Y: 0.036**	**Y: 0.034**	**Y: 0.037**	**Y: 0.031**	0.026
Sensitivity	Single SVM	**Y: 0.182**	N: 0.167	N: 0.168	N: 0.168	**Y: 0.220**	N: 0.124	0.176
	SVM ensemble	**Y: 0.114**	**Y: 0.094**	**Y: 0.116**	**Y: 0.121**	**Y: 0.113**	N: 0.084	0.092
	Single MDA	**Y: 0.172**	**Y: 0.133**	**Y: 0.167**	N: 0.128	**Y: 0.168**	N: 0.101	0.132
	MDA ensemble	**Y: 0.147**	**Y: 0.121**	**Y: 0.123**	**Y: 0.118**	**Y: 0.127**	N: 0.091	0.118

Measure	Method	No selection	B&B5	B&B10	Factor analysis	t-test	t-test & CA	BBHFS
Specificity	Single SVM	**Y:** 0.021	N: 0.017	N: 0.018	**Y:** 0.019	**Y:** 0.025	**Y:** 0.025	0.019
	SVM ensemble	**Y:** 0.054	**Y:** 0.032	**Y:** 0.037	**Y:** 0.048	**Y:** 0.061	**Y:** 0.048	0.031
	Single MDA	**Y:** 0.028	**Y:** 0.017	**Y:** 0.020	**Y:** 0.022	**Y:** 0.023	**Y:** 0.019	0.016
	MDA ensemble	**Y:** 0.048	N: 0.031	**Y:** 0.041	**Y:** 0.042	**Y:** 0.045	**Y:** 0.038	0.032
Number of **Y**		12	8	9	9	12	7	

4.2.3. Comparison between IOMCE and single-classifier models

Based on Table 9, we calculate the difference between specificity and sensitivity for the FDP data set and show the results in Table 12. Table 13 shows the differences in accuracy, sensitivity, and specificity between IOMCE and single-classifier models for the FDP data set. For the visualization of the relationships between single-classifier and IOMCE models in terms of their performance evaluation measures, we provide graphs in Figure 4 based on the results in Table 9.

Table 12. Differences between specificity and sensitivity for the FDP data set

Method	No selection	B&B5	B&B10	Factor analysis	t-test	t-test & CA	BBHFS
Single SVM	40.71	38.48	42.63	39.94	50.26	42.03	40.09
SVM ensemble	8.52	9.37	9.97	9.52	2.57	10.61	9.70
Single MDA	30.85	26.94	30.80	25.46	30.90	30.01	26.43
MDA ensemble	12.58	14.13	10.95	10.39	8.30	8.63	11.96

Table 13. Differences in accuracy, sensitivity, and specificity between IOMCE and single-classifier models for the FDP data set

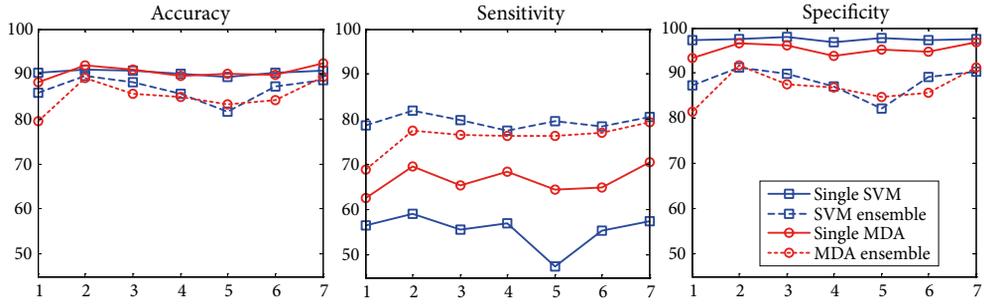| Measure	Method	No selection	B&B5	B&B10	Factor analysis	t-test	t-test & CA	BBHFS
Accuracy	SVM	–4.62	–1.46	–2.77	–4.64	–7.56	–2.83	–2.12
	MDA	–8.75	–2.76	–5.32	–4.55	–6.71	–5.46	–3.04
Sensitivity	SVM	22.09	22.68	24.32	20.59	32.00	23.23	23.09
	MDA	6.41	7.86	11.14	7.95	12.04	12.28	8.96
Specificity	SVM	–10.10	–6.43	–8.34	–9.83	–15.69	–8.19	–7.30
	MDA	–11.86	–4.95	–8.71	–7.12	–10.56	–9.10	–5.51

Fig. 4. Graphs of performance evaluation measures of single-classifier and IOMCE models
for the FDP data set

Consistent with the results for the corporate credit risk data set, the differences in specificity and sensitivity are much larger for single-classifier models than for IOMCE models (Table 12), and the reductions in accuracy and specificity are generally smaller than the increases in sensitivity except for MDA with no feature selection (Table 13). In addition, the relative locations of the curves in Figure 4 are very similar to those in Figure 3, providing further support for the argument that the IOMCE method can dramatically improve the recognition rates for the minority group while incurring only small reductions in overall accuracy and specificity. These results and those in Section 4.2.2 together suggest that the BBHFS-IOMCE method can still produce satisfactory performance for a more imbalanced (1:5) data set with more initial features (42). It is also indicated that processing an imbalanced data set with the IOMCE method is even more necessary for the SVM than for the MDA, which is consistent with the results for the corporate credit risk data set. This phenomenon may be due to the fact that the decision function of the SVM is determined only by support vectors, whereas that of the MDA is determined by all samples, which brings more effect of data imbalance on the SVM than on the MDA.

4.3. Experimental results and analysis for the personal credit risk data set

The results for different feature selection methods for the personal credit risk data set indicate no significant superiority of BBHFS over the others and no significant advantages of the four feature selection methods over no feature selection, which is inconsistent with the results for the corporate credit risk data set and the FDP data set. This may be due to the fact that the personal credit risk data set has much more discrete and nominal features than continuous ones and that this study's feature selection methods are more suitable for the latter. For example, Dash and Liu (1997) indicate that nominal features require special handling for feature selection because it is not easy to assign real values to them and B&B is not applicable to such feature selection. In addition, the t-test may not be valid when discrete data sets are analyzed (McElduff *et al.* 2010), and the factor analysis may not be appropriate for non-orderable discrete indicators with more than two categories (Ender 2005). In this regard, we only provide the results for no feature selection for the personal credit risk data set in Table 14 and chart mean accuracy, sensitivity, and specificity for different models in Figure 5.

Table 14. Comparison of mean accuracy, sensitivity, and specificity between single-classifier and IOMCE models for the personal credit risk data set

Measure	Method	Single classifier	IOMCE	Difference	Increase/decrease (%)
Accuracy	SVM	78.43	70.39	−8.04	−10.25%
	MDA	71.71	71.04	−0.67	−0.93%
Sensitivity	SVM	37.45	69.37	31.92	85.23%
	MDA	70.77	72.68	1.91	2.70%
Specificity	SVM	92.09	70.73	−21.36	−23.19%
	MDA	72.03	70.49	−1.54	−2.14%
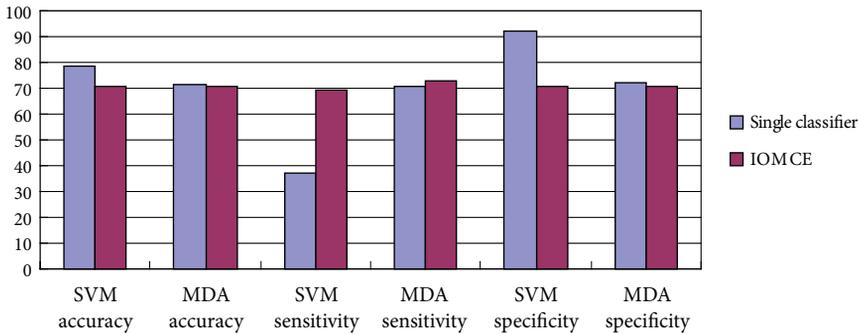
Fig. 5. Graphs of performance evaluation measures for single-classifier and IOMCE models for the person credit risk data set

The results indicate that the IOMCE method can improve the recognition rates for minority samples while incurring acceptable reductions in accuracy and specificity when the data set contains many discrete and nominal features. In addition, the role of the IOMCE method in processing an imbalanced data set is more evident for the SVM than for the MDA, which is consistent with the results for the corporate credit risk data set and the FDP data set. This suggests that the MDA is more robust than the SVM when a single-classifier model is used for an imbalanced data set.

Conclusions

This paper proposes an integrated credit risk assessment method that combines BBHFS with the IOMCE and uses the SVM and the MDA as the base predictor. We use BBHFS to select a feature subset with an optimal feature extraction rate and the IOMCE to address an imbalanced data set. For comparative analysis, we also used no feature selection and three other feature selection methods (pure B&B, the factor analysis, and the pure t-test) in single-classifier or IOMCE modeling. We compared their results with those of single-classifier and IOMCE models based on BBHFS. The empirical results for the corporate credit risk data set and two extended data sets show that credit risk modeling based on the BBHFS-IOMCE method generally produces the best classification performance when all features are continuous.

BBHFS is able to retain useful information for credit risk assessment when reducing the dimensionality of continuous features, but it is not suitable for discrete and nominal features. With either continuous or discrete/nominal features, using the IOMCE method for an imbalanced data set can significantly reduce the differences between specificity and sensitivity by increasing sensitivity sharply and reducing specificity only slightly. Taken together, these results suggest that both BBHFS and the IOMCE are necessary for credit risk modeling and that the BBHFS-IOMCE method can produce competitive performance for imbalanced data sets with continuous features.

Acknowledgements

References

Ahn, H.; Moon, H.; Fazzari, M.; Lima, N.; Chenb, J. J.; Kodell, R. L. 2007. Classification by ensembles from random partitions of high-dimensional data, *Computational Statistics & Data Analysis* 51(6): 166–177.

Angelini, E.; Tollo, G.; Roli, A. 2008. A neural network approach for credit risk evaluation, *The Quarterly Review of Economics and Finance* 48(4): 733–755. http://dx.doi.org/10.1016/j.qref.2007.04.001

Baesens, B.; Setiono, R.; Mues, C.; Vanthienen, J. 2003. Using neural network rule extraction and decision tables for credit risk evaluation, *Management Science* 49(3): 312–329. http://dx.doi.org/10.1287/mnsc.49.3.312.12739

Bellotti, T.; Crook, J. 2008. Support vector machines for credit scoring and discovery of significant features, *Expert Systems with Applications* 36(2): 3302–3308. http://dx.doi.org/10.1016/j.eswa.2008.01.005

Bellotti, T.; Crook, J. 2009. Credit scoring with macroeconomic variables using survival analysis, *Journal of the Operational Research Society* 60(12): 1699–1707. http://dx.doi.org/10.1057/jors.2008.130

Bellotti, T.; Matousek, R.; Stewart, C. 2011. A note comparing support vector machines and ordered choice models' predictions of international banks' ratings, *Decision Support Systems* 51(3): 682–687. http://dx.doi.org/10.1016/j.dss.2011.03.008

Besanko, D.; Thakor, A. 1987. Collateral and rationing: sorting equilibria in monopolistic and competitive credit markets, *International Economic Review* 28(3): 671–689. http://dx.doi.org/10.2307/2526573

Bester, H.; Sreening, V. 1985. Rationing in credit markets with imperfect information, *American Economics Review* 75(4): 850–855.

Boguslauskas, V.; Mileris, R.; Adlyte, R. 2011. New internal rating approach for credit risk assessment, *Technological and Economic Development of Economy* 17(2): 369–381. http://dx.doi.org/10.3846/20294913.2011.583721

Brown, I.; Mues, C. 2012. An experimental comparison of classification algorithms for imbalanced credit scoring data sets, *Expert Systems with Applications* 39(3): 3446–3453.

Chawla, N. V.; Bowyer, K. W.; Hall, L. O.; Kegelmeyer, W. P. 2002. SMOTE: synthetic minority over-sampling technique, *Journal of Artificial Intelligence Research* 16: 321–357.

Chawla, N. V.; Japkowicz, N.; Kolcz, A. 2004. Editorial: special issue on learning from imbalanced data sets, *SIGKDD Explorations* 6(1): 1–6. http://dx.doi.org/10.1145/1007730.1007733

Chen, X. 2003. An improved branch and bound algorithm for feature selection, *Pattern Recognition Letters* 24(12): 1925–1933. http://dx.doi.org/10.1016/S0167-8655(03)00020-5

Chen, C.; Liaw, A.; Breiman, L. 2004. *Using random forests to learn unbalanced data.* California: University of California. 12 p.

Chen, W.; Ma, C.; Ma, L. 2007. Mining the customer credit using hybrid support vector machine technique, *Expert Systems with Applications* 36(4): 7611–7616. http://dx.doi.org/10.1016/j.eswa.2008.09.054

Chijoriga, M. M. 2011. Application of multiple discriminant analysis (MDA) as a credit scoring and risk assessment model, *International Journal of Emerging Markets* 6(2): 132–147.

Cimpoeru, S. 2011. Neural networks and their application in credit risk assessment. Evidence from the Romanian market, *Technological and Economic Development of Economy* 17(3): 519–534. http://dx.doi.org/10.3846/20294913.2011.606339

Constantinopoulos, C.; Likas, A. 2008. Semi-supervised and active learning with the probabilistic RBF classifier, *Artificial Neural Networks* 71(13): 2489–2498. http://dx.doi.org/10.1016/j.neucom.2007.11.039

Crook, J.; Edelman, D.; Thomas, L. 2007. Recent developments in consumer credit risk assessment, *European Journal of Operational Research* 183(3): 1447–1465. http://dx.doi.org/10.1016/j.ejor.2006.09.100

Dash, M.; Liu, H. 1997. Feature selection for classification, *Intelligent Data Analysis* 1(1–4): 131–156. http://dx.doi.org/10.1016/S1088-467X(97)00008-5

Dong, G.; Lai, K.; Yen, J. 2010. Credit scorecard based on logistic regression with random coefficients, *Procedia Computer Science* 1(1): 2463–2468. http://dx.doi.org/10.1016/j.procs.2010.04.278

Eisenbeis, R. 1997. Pitfalls in the application of discriminant analysis in business, finance, and economics, *Journal of Finance* 32(3): 875–900. http://dx.doi.org/10.1111/j.1540-6261.1977.tb01995.x

Ender, P. 2005. *Multivariate analysis: more exploratory factor analysis* [online], [cited 07 May 2012]. Available from Internet: http://www.philender.com/courses/multivariate/notes2/morefa.html

Fan, R.-E.; Chen, P.-H.; Lin, C.-J. 2005. Working set selection using second order information for training SVM, *Journal of Machine Learning Research* 6: 1889–1918.

Finlay, S. 2008. Towards profitability: a utility approach to the credit scoring problem, *Journal of the Operational Research Society* 59(7): 921–931. http://dx.doi.org/10.1057/palgrave.jors.2602394

Fisher, R. A. 1936. The use of multiple measurements in taxonomic problems, *Annals of Eugenics* 7(2): 179–188. http://dx.doi.org/10.1111/j.1469-1809.1936.tb02137.x

Florez-Lopez, R. 2010. Effects of missing data in credit risk scoring. A comparative analysis of methods to achieve robustness in the absence of sufficient data, *Journal of the Operational Research Society* 61(3): 486–501. http://dx.doi.org/10.1057/jors.2009.66

Garcia, V.; Sanchez, J. S.; Mollineda, R. A. 2012. On the effectiveness of preprocessing methods when dealing with different levels of class imbalance, *Knowledge-Based Systems* 25(1):13–21.

Guyon, I.; Elisseeff, A. 2003. An introduction to variable and feature selection, *Journal of Machine Learning Research* 3: 1157–1182.

Heijden, F.; Duin, R.; Ridder, D.; Tax, D. 2004. *Classification, parameter estimation and state estimation: an engineering approach using MATLAB*. The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England: Wiley. 423 p. http://dx.doi.org/10.1002/0470090154

Hillier, B.; Lbrahimo, M. 1993. Asymmetric information and models of credit rationing, *Bulletin of Economic Research* 45(4): 271–304. http://dx.doi.org/10.1111/j.1467-8586.1993.tb00571.x

Hofmann, H.; Stat, L. 1994. *UCI machine learning repository* [online], [cited 07 May 2012]. Available from Internet: http://archive.ics.uci.edu/ml

Huang, S. 2009. Integrating nonlinear graph based dimensionality reduction schemes with SVMs for credit rating forecasting, *Expert Systems with Applications* 36(4): 7515–7518. http://dx.doi.org/10.1016/j.eswa.2008.09.047

Huang, Z.; Chen, H.; Hsu, C.-J.; Chen, W.-H.; Wu, S. 2004. Credit rating analysis with support vector machine and neural networks: a market comparative study, *Decision Support Systems* 37(4): 543–558. http://dx.doi.org/10.1016/S0167-9236(03)00086-1

Hui, X.; Sun, J. 2006. An application of support vector machine to companies' financial distress prediction, *Lecture Notes in Computer Science* 3885: 274–282. http://dx.doi.org/10.1007/11681960_27

Kennedy, K.; Namee, B. M.; Delany, S. J. 2009. Learning without default: a study of one-class classification and the low-default portfolio problem, in *Proc. of the 20th Irish Conference on Artificial Intelligence and Cognitive Science*, 19–21 August 2009, Dublin, 174–187.

Khashman, A. 2010. Neural networks for credit risk evaluation: investigation of different neural models and learning schemes, *Expert Systems with Applications* 37: 6233–6239.

Kubat, M.; Matwin, S. 1997. Addressing the curse of imbalanced training sets: one-sided selection, in *Proceedings of the 14th International Conference on Machine Learning*, 8–12 July 1997. San Francisco: Morgan Kaufmann, 179–186.

Laha, A. 2007. Building contextual classifiers by integrating fuzzy rule based classification technique and k-nn method for credit scoring, *Advanced Engineering Informatics* 21: 281–291.

Lahsasna, A.; Ainon, R.; Wah, T. 2010. Credit scoring models using soft computing methods: a survey, *International Arab Journal of Information Technology* 7(2): 115–123.

Laitinen, E. 1999. Predicting a corporate credit analysis risk estimate by logistic and linear models, *International Review of Financial Analysis* 8(2): 97–121. http://dx.doi.org/10.1016/S1057-5219(99)00012-5

Lawler, E.; Wood, D. 1966. Branch-and-bound method: a survey, *Operations Research* 14(4): 699–719. http://dx.doi.org/10.1287/opre.14.4.699

Liu, J.; Li, J.; Xu, W. 2011. A weighted L(q) adaptive least squares support vector machine classifiers – robust and sparse approximation, *Expert Systems with Applications* 38(3): 2253–2259.

Liu, Y.; Wu, C.; Wang, M.; Qiao, M. 2005. The model of credit risk assessment in commercial banks on support vector machine, *Forecasting* 1(24): 52–55.

Maloof, M. 2003. Learning when data sets are imbalanced and when costs are unequal and unknown, in *ICML-2003 Workshop on Learning from Imbalanced Data Sets II*, 21–24 August 2003, Washington DC: AAAI Press, 1–8.

Martens, D.; Baesens, B.; Gestel, T.; Vanthienen, J. 2007. Comprehensible credit scoring models using rule extraction from support vector machines, *European Journal of Operational Research* 183(3): 1466–1476. http://dx.doi.org/10.1016/j.ejor.2006.04.051

Martens, D.; Gestel, T.; Backer, M.; Raf, H.; Jan, V.; Christophe, M.; Bart, B. 2010. Credit rating prediction using ant colony optimization, *Journal of the Operational Research Society* 61: 561–573. http://dx.doi.org/10.1057/jors.2008.164

McElduff, F.; Cortina-Borja, M.; Chan, S.-K.; Wade, A. 2010. When t-tests or Wilcoxon-Mann-Whitney tests won't do, *Advances in Physiology Education* 34(3): 128–133. http://dx.doi.org/10.1152/advan.00017.2010

Paleologo, G.; Elisseeff, A.; Antonini, G. 2010. Subagging for credit scoring models, *European Journal of Operational Research* 201(2): 490–499. http://dx.doi.org/10.1016/j.ejor.2009.03.008

Spearman, C. 1904. General intelligence, objectively determined and measured, *The American Journal of Psychology* 15(2): 201–292. http://dx.doi.org/10.2307/1412107

Stiglitz, J.; Wesiss, A. 1981. Credit rationing in markets with imperfect information, *American Economics Review* 71(3): 393–410.

Sun, J.; Li, H. 2011. Dynamic financial distress prediction using instance selection for the disposal of concept drift, *Expert Systems with Applications* 38(3): 2566–2576.
http://dx.doi.org/10.1016/j.eswa.2010.08.046

Tsai, C. 2009. Feature selection in bankruptcy prediction, *Knowledge-Based Systems* 22(2): 120–127.
http://dx.doi.org/10.1016/j.knosys.2008.08.002

Tsai, C.; Wu, J. 2008. Using neural network ensembles for bankruptcy prediction and credit scoring, *Expert Systems with Applications* 34(4): 2639–2649. http://dx.doi.org/10.1016/j.eswa.2007.05.019

Tseng, H.; Lin, G.; Tsai, L. 2011. An enforced support vector machine model for construction contractor default prediction, *Automation in Construction* 20(8): 1242–1249.
http://dx.doi.org/10.1016/j.autcon.2011.05.007

Vapnik, V. 1996. *The nature of statistical learning theory*. New York: Springer-Verlag. 188 p.

Wang, G.; Hao, J.; Ma, J.; Jiang, H. 2011. A comparative assessment of ensemble learning for credit scoring, *Expert Systems with Applications* 38(1): 223–230. http://dx.doi.org/10.1016/j.eswa.2010.06.048

West, D.; Dellana, S.; Qian, J. 2005. Neural network ensemble strategies for financial decision applications, *Computers & Operations Research* 32(10): 2543–2559. http://dx.doi.org/10.1016/j.cor.2004.03.017

Williamson, S. 1986. Costly monitoring, financial intermediation, and equilibrium credit rationing, *Journal of Monetary Economics* 18(2): 159–179. http://dx.doi.org/10.1016/0304-3932(86)90074-7

Wu, C.; Guo, Y.; Xia, H. 2009. The model of credit risk assessment in commercial banks on fuzzy integral support vector machines ensemble, *Operations Research and Management Science* 2(18): 115–119.

Wu, C.; Guo, Y.; Zhang, X.; Xia, H. 2010. Study of personal credit risk assessment based on support vector machine ensemble, *International Journal of Innovative Computing Information and Control* 6(5): 2353–2360.

Xie, C.; Luo, C.; Yu, X. 2011. Financial distress prediction based on SVM and MDA methods: the case of Chinese listed companies, *Quality & Quantity* 45(3): 671–686.
http://dx.doi.org/10.1007/s11135-010-9376-y

Xu, X.; Chen, Y.; Zheng, H. 2011. The comparison of enterprise bankruptcy forecasting method, *Journal of Applied Statistics* 38(2): 301–308. http://dx.doi.org/10.1080/02664760903406470

Yan, R.; Liu, Y.; Jin, R.; Alex, H. 2003. On predicting rare classes with SVM ensembles in scene classification, in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 6–10 April 2003, Hong Kong, 21–24.

Yei, Z.; Wen, Y.; Lu, B. 2009. A survey of imbalanced pattern classification problems, *CAAI Transactions on Intelligent Systems* 4(2): 148–153.

Yu, L.; Wang, S.; Lai, K. 2008. Credit risk assessment with a multistage neural network ensemble learning approach, *Expert Systems with Applications* 34(2): 1434–1444.
http://dx.doi.org/10.1016/j.eswa.2007.01.009

Yu, L.; Wang, S.; Lai, K. 2009. An intelligent-agent-based fuzzy group decision making model for financial multi-criteria decision support: the case of credit scoring, *European Journal of Operational Research* 195(3): 942–959. http://dx.doi.org/10.1016/j.ejor.2007.11.025

Yu, L.; Yao, X.; Wang, S.; Lai, K. K. 2011. Credit risk evaluation using a weighted least squares SVM classifier with design of experiment for parameter selection, *Expert Systems with Applications* 38(12): 15392–15399. http://dx.doi.org/10.1016/j.eswa.2011.06.023

Yu, L.; Yue, W.; Wang, S.; Lai, K. 2010. Support vector machine based multi-agent ensemble learning for credit risk evaluation, *Expert Systems with Applications* 37: 1351–1360.

Zhou, L.; Lai, K.; Yu, L. 2010a. Least squares support vector machines ensemble models for credit scoring, *Expert Systems with Applications* 37(1): 127–133. http://dx.doi.org/10.1016/j.eswa.2009.05.024

Zhou, X.; Jiang, W.; Shi, Y. 2010b. Credit risk evaluation by using nearest subspace method, *Procedia Computer Science* 1(1): 2449–2455. http://dx.doi.org/10.1016/j.procs.2010.04.276

Jie SUN received her BS, MS, and PhD degrees all with honor from Harbin Institute of Technology (HIT), Harbin, China. The author's major field of study includes financial risk management, statistical and intelligent modeling for economics, concept drift and financial intelligence. She is Professor of Zhejiang Normal University, Jinhua, Zhejiang, China, and was a Guest Professor of Dongguk University, Gyeongju, Korea. She serves as Principal Investigator of several national funded research projects including the National Natural Science Foundation of China, the Zhejiang Provincial Natural Science Foundation of China, and Zhejiang Provincial Philosophy and Social Science Foundation of China. She received the Outstanding MS Dissertation Award of HIT in 2004, the Outstanding PhD Dissertation Award of HIT in 2008, the Science Research Awards of Zhejiang Provincial Universities in 2010 and 2011, and the Social Science Research Awards of Jinhua City in 2010 and 2012. She is listed in the Marquis Who's Who in the World since 2010. Her current research interest is concept drift oriented support vector machine ensemble in financial crisis management.

Young-Chan LEE is an Associate Professor of Department of Information Management and Dean of International Affairs at Dongguk University, GyeongJu, Korea. He also serves as Editor-in-Chief at Korea Association of Information Systems. He received a PhD degree of Management Science in Department of Business Administration from Sogang University, Seoul, Korea. He has published papers on some international journals. He also has published tens of books, and won a lot of research funding and participated in various projects. He was awarded the "Most Cited Papers" in Expert Systems with Applications in 2006–2008, and is listed in the Marquis Who's Who in the World since 2007. He was awarded as Best Paper of Korea Association of Information Systems in 2009 and 2010 successively. His research interests include business intelligence, enterprise systems, green information system and sustainability, and multi-criteria decision making.

Hui LI received his BS, MS and PhD degrees all with honor from Harbin Institute of Technology, Harbin, Heilongjiang, China. The author's major field of study includes case-based reasoning, business forecasting, business computing, business data mining, and business intelligence, among others. He is currently an Associate Professor of Zhejiang Normal University, Jinhua, Zhejiang, China and was a Visiting Scholar of The Ohio State University, Columbus, OH, US. He is a young Researcher of World Federation on Soft Computing, a Member of Association for Information Systems, a Member of the Professional Risk Managers' International Association, and a Zhijiang Young Talent in Social Science. He led 7 national and provincial funds, and serves as International Committee Member of several international conferences, Editorial Board Member of several international journals, and referees of many leading journals. He received the Outstanding MS Dissertation Award of HIT in 2004, the Outstanding MS Dissertation Award of Heilongjiang Province in 2006, the Outstanding PhD Dissertation Award of HIT in 2009, the Award of Outstanding Young Talents of Zhejiang Province in 2009, the Science Research Awards of Zhejiang Provincial Universities in 2010 and 2011, the Social Science Research Awards of Jinhua in 2010 and 2012, and is listed on Marquis Who's Who in the World since 2010.

Qing-Hua HUANG received her both BS and MS degrees from Zhejiang Normal University, Jinhua, Zhejiang, China. She received the fostering funding of outstanding MS dissertation from Zhejiang Normal University. The author's major study field includes: financial risk prediction, accounting, and imbalanced modeling in business.