# Kybernetes

Evaluating credit rating prediction by using the KMV model and random forest

Hsu-Che Wu Yu-Ting Wu

# Evaluating Credit Rating Prediction by Using the KMV Model and Random Forest



## 1. INTRODUCTION

After misrepresenting financial data, the Enron Corporation and WorldCom filed bankruptcy in December 2001 and July 2002, respectively. Subsequently, the subprime mortgage crisis occurred in July 2007. In July 2008, Freddie Mac and Fannie Mae experienced a bankruptcy crisis, and in September 2008, the severe disruptions caused by the subprime mortgage crisis led to a severe global recession. Consequently, Merrill Lynch was sold to the Bank of America, Lehman Brothers Holdings Inc. filed bankruptcy, and the American International Group underwent a financial crisis. In 2011, the United States Senate accused the three largest credit rating agencies, Standard & Poor's, Moody's Investors Service, and Fitch Ratings, of providing overrated promissory notes and questioned the credit rating standards among these companies. In other words, their power may have affected the market. These incidents implied the lack of transparent credit rating information resulting from managerial misconduct.Thus, novel credit rating methods should play a crucial role in the capital market.

Previously, long-term credit risk classification focused on forecasting credit ratings. The models proposed by Black and Myron (1973) and Merton (1974) are the most prominent examples. These models involve using option pricing theory, stockholders' equity market value, and stockholders' equity value for estimating the credit risk premium of corporate bonds. Since the late 1980s, financial and accounting data have been used to develop statistical or financial models for estimating the distance to default (DD). In an early example of financial failure prediction, Beaver (1966) used a univariate analysis as a predictive model. Researchers have subsequently leveraged artificial intelligence to generate novel credit rating methods such as the neural networks (NN), genetic algorithm (GA), decision tree (DT), and support vector machine (SVM). These classification methods differ from conventional regression analysis models, yielding superior classification. Breiman (2001) proposed the random forest (RF) and used the classification and regression tree to perform feature selection. Díaz-Uriarte and Andrés (2006) were the first to use the RF to perform feature selection, implementing an advanced DT technique for use in gene data selection; they were also first to upgrade the classification method to a multivariate method. Strobl et al. (2008) proved that the RF is a useful tool for selecting critical variables from a group of highly correlated variables. In addition, the RF has been applied to credit rating research (e.g., Mori & Umezawa, 2007; Yeh et al., 2012) to analyze credit rating indicators, and the results have suggested that RF-based credit rating models were highly accurate. However, the credit rating service in the current capital market is monopolized by few the credit rating agencies. The rating system is excessively complex for investors to use, credit rating agencies charge exorbitant service fees, and the rating methods are not sufficiently transparent. New methods are constantly being researched; however, these methods (Table 2) provide weak feature selection and explanation of variables.

In this study, a credit rating model was used to calculate the estimated DD, which was applied in a consideration framework to determine whether the forecasting ability of the credit rating model was enhanced. This information can serve as a reference for investors and potentially reduce credit risk in the banking industry. In addition, this study proposes using an RF model and data-mining techniques to construct a two-step credit rating model. Various tools for evaluating key credit rating indicators, which were derived from the financial and nonfinancial variables discussed in relevant literature, were applied for enhancing the credit rating system and thereby improving decision making; the DD obtained from the Kealhofer, McQuown, and Vasicek (KMV) model was used to form a novel model. In addition, Standard & Poor's CompuStat Research Insight dataset, which was derived from a financial analysis of public companies in North America, was employed and combined with the RF, DT, NN, or SVM models to develop a sufficiently accurate credit rating model. The accuracy levels of several credit rating models were compared in order to identify the optimal model. The results of this comparison can be referenced in internal risk management and should provide a valuable reference for external investors. This study features the following research methods:

- Using the RF to perform feature selection and determine whether this addition enhances the forecasting capacity of the credit rating model.
- Comparing the levels of forecast efficiency among several classification models.
- Creating a two-step credit rating model by using data-mining techniques.



## 2. LITERATURE REVIEW

### 2.1 The Definition and Development of Credit Rating

The credit rating system was established in the early 20th century in the United States. John Moody founded Moody's Investors Service, and when he first graded the railroad bond rating during the Great Depression, one-third of bonds were at risk of default. Investors gradually realized the importance of the credit rating system, which developed to encompass all types of financial products and evaluator applications.

Credit rating agencies perform credit ratings to evaluate the value and default rate of an enterprise; transparent and comprehensive rating methods and standards are used to investigate and analyze public financial ratios, thereby characterizing the financial condition of a case company. Initial letter or credit level numbers are used to differentiate the enterprise risk indicators. Credit ratings are used to evaluate the credit risk of debtors in the capital market. The Association of Banks uses credit rating as a form of assessment. A credit score chart is generated and quantitative methods are used to grade and comment on the credit factors affecting debtors, including their financial conditions, management, industry characteristics, and prospects. The weight is then determined using a concrete method and the credit rating level is yielded according to the score earned. Mays (2001), Mays and Mays (1998), and Bailey et al. (2004) have described the establishment of credit score cards. Recent developments in the credit rating system have been focused on using the financial ratio to determine whether enterprise management bases on better performance. Previous research on enterprise crises has predominantly considered the possibility or forecasting of financial crises. Gupton et al. (1997) determined that credit risk forecasting models can be used to predict the credit default rate and losses following default. Kliger and Sarig (2000) reported that announcing a credit rating can effectively reflect the enterprise value in the capital market. Such an announcement eliminates information asymmetry, enabling investors to reasonably evaluate security prices and reducing the possibility of abnormal returns.

## 2.2 The KMV Model

KMV was founded in 1989. The company applied Merton’s (1974) option pricing model to evaluate the default rate of an enterprise and used Black and Myron’s (1973) option pricing theory to evaluate its predicted default rate. KMV modified Merton’s (1974) model to evaluate enterprise assets and observe default conditions, using data from a database to calculate the DD of an enterprise. When the assets of an enterprise fall below the default point, the enterprise breaks its contract. In the KMV model, the asset volatility of a market-information-oriented enterprise is the key indicator for evaluating the default rate. In this study, the DD was added as a key indicator for evaluating enterprise assets and the default rate. Table 1 lists a summary of studies using the DD from the KMV model as a research variable.

**Table 1 Relevant research on the KMV model**

| Author | Research Findings |
|---|---|
| **Lopez (2004)** | The KMV model and its asymptotic single risk factor were used to analyze risk capital. Furthermore, the initial book value of the assets was studied, and the correlations among average asset, default rate, and asset scale were evaluated. |

| Author | Research Findings |
|---|---|
| **Du and Suo (2007)** | A multiple regression model was created and variables from the KMV model were added as dependent variables used to forecast the credit rating of enterprises undergoing a financial crisis. |
| **Bharath and Shumway (2008)** | Three types of hypothesis were proposed to inspect the calculated DD obtained using the KMV model and the DD was compared with the Z-score model. |
| **Campbell et al. (2008)** | A multiple regression model was used to combine financial variables and DD variables obtained using the KMV model to determine whether the enterprise earnings and low stock price performance were affected by a financial crisis. |
| **Yeh et al. (2012)** | The RF model was used to create a credit rating model, which was classified using a rough set to satisfy the decision rule. The model was then combined with the KMV model and corporate governance variables to improve credit rating evaluation. |

### 2.3 Applying Data Mining to Credit Ratings

In previous studies, univariate analysis was most frequently applied in forecasting models.Researchers have typically used financial variables to predict financial crises. Beaver (1966) first used univariate analysis to develop an early financial warning model; financial ratios could yield effective predictions when conducting univariate analysis. Univariate analysis can be used to explain the financial condition of an enterprise before a financial crisis occurs based on a single variable. This method is advantageous because it enables simple calculations.This method became a fundamental model for generating credit rating models; however, in current, complex markets, using only single variable is ineffective. In addition, classification results that yield distinct ratios might cause antinomy. In practice, using only a single financial ratio to evaluate the performance level of an enterprise is difficult. The situation is exacerbated when evaluating several enterprises. As shown in Table 2, the references regarding data-mining classification techniques were organized according to year, indicating that enhancing the efficiency of credit rating forecasts is critical. The results listed in Table 2 indicate that superior forecasting efficiency can be achieved using a suitable data mining technique.

**Table 2 Relevant research on data mining techniques applied to credit rating**

| Authors | Method | Research Findings |
|---|---|---|

| | | |
|---|---|---|
| **Dutta and Shekhar (1988)** | NN | Dutta and Shekhar used an NN in bond credit rating in the early study. The results indicated that the accuracy was between 76%–82%, suggesting that the NN exhibited a performance superior to that of regression analysis models. |
| **Odom and Sharda (1990)** | NN<br>Multivariate Analysis | Odom and Sharda employed data-mining techniques and statistical models to create credit rating models and compared several sampling proportions to improve linear models. The results indicated that the new model exhibited a more favorable forecasting performance than the linear models did. |
| **Altman and Saunders (1997)** | NN | Altman and Saunders applied an NN in creating a credit rating model, subsequently compared the model with that of the aforementioned study, and conducted discriminant analysis. The results indicated that both models achieved 90% accuracy. |
| **Maher and Sen (1997)** | NN | Maher and Sen used a back propagation neural network (BPN) and multiple regression analysis to create a credit rating model, and compared its accuracy. The results indicated that the classification accuracy of both models achieved 70% accuracy. |
| **Shin and Han (2001)** | Case-Based Reasoning | Case-based reasoning was used to construct Korean credit rating models. Shin and Han employed 12 variables to classify 3,886 enterprises into five levels. The forecasting capacity was 70%. |
| **Chen and Shih (2006))** | SVM | Chen and Shih applied the SVM with financial variables to evaluate the credit rating of the banking industry and compared the SVM with traditional models to determine the accuracy rate. The results indicated that SVM exhibited a performance 84% higher than that of the BPN and regression model. |
| **Zhou and Bai (2008)** | SVM<br>Genetic Algorithms<br>Rough Set | Genetic algorithms and a rough set were combined to select and evaluate variables, and then their prediction accuracy was compared. The results indicated that using the rough set in initial variable selection increased prediction accuracy. |
| **Angelini et al. (2008)** | NN | An NN was applied to create a credit rating model by using15 financial variables to classify 78 samples of Italian companies. The accuracy was 93%. |

| | | |
|---|---|---|
| **Yeh et al. (2010)** | Data Envelopment Analysis<br>Rough Set<br>RF | Yeh et al. employed a rough set and the RF to create a credit rating model featuring high prediction accuracy, and then added DEA to compare the operating efficiency among variables used in forecasting. The results indicated that after DEA was added, the forecasting accuracy was 88%; when only financial data were added, the forecasting accuracy was 76%, revealing that two-step forecasting exhibits superior performance. |
| **Härdle et al. (2012)** | SVM<br>BPN<br>Rough Set<br>Fuzzy Theory | Härdle et al. combined rough set and fuzzy theory to perform variable selection, and then added the SVM to determine the optimal hyperplane before create credit rating classification. The results indicated that the rough set yielded the highest forecasting rate, 89.42%, followed by SVM, which yielded an 87.72% forecasting rate, and then BPN, which yielded an 81.14% forecasting rate. |
| **Chen and Cheng (2013)** | Rough Set<br>RF<br>DT | Chen and Cheng integrated the KMV model and DD with financial variables to analyze credit rating by using the RF. The results indicated that the combination of a rough set and the RF yielded the highest forecasting rate of 93.4%. |

## 3. Research Method

### 3.1Experiment Design

This study involved constructing a mixed model based on data-mining techniques. An RF model was applied to facilitate selecting the critical variables and filtering the repeated and nonrelated variables. Most data exhibiting one or several nonrelated attributes was eliminated, and valid features of the data were extracted (Díaz-Uriarte &Andrés, 2006). Subsequently, these featured variables and different data-mining classification techniques were used to evaluate credit rating models; this enabled exploring the prediction efficiency among the models.

The R language was used as a feature selection tool; this language provides several statistical methods (e.g., linear models, nonlinear models, time-series analysis, and clustering) and is highly adaptable and extendable. Furthermore, the RF, DT, NN, and SVM models were used to create credit rating classifiers based on the Waikato Environment for Knowledge Analysis (WEKA), and the forecasting performance level of each model was compared.

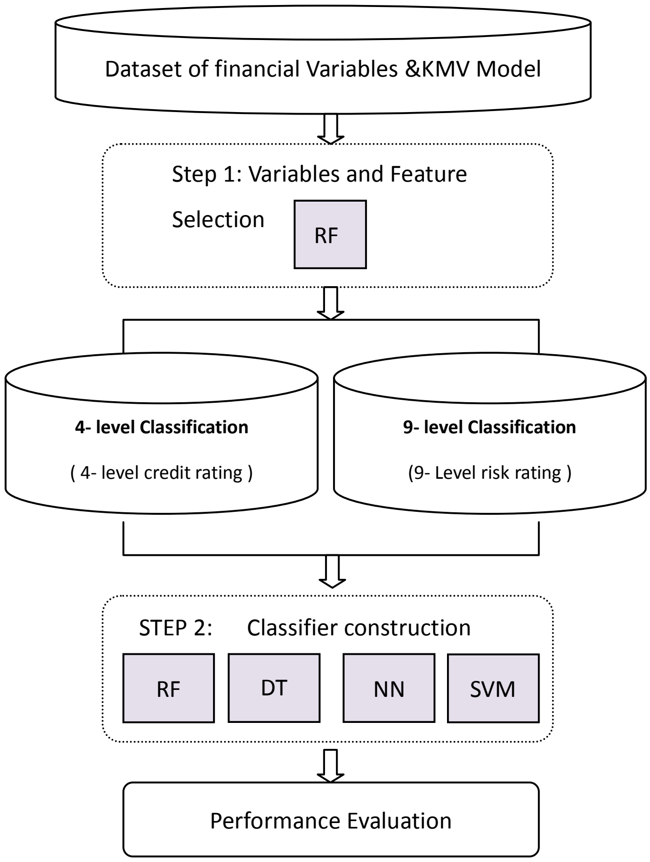


**Figure 1 The research process**

### 3.2 Goals of Classification

In 1922, Standard & Poor's began developing bond-issue ratings, involving industry risk and management, enterprise strategy, sales evaluation, management evaluation, investment, the capitalization processes, liquidity, and financial flexibility. The financial strength rating was a key indicator when evaluating whether the general financial condition of an enterprise was adequate for fulfilling obligations to debtors. Ratings can be classified into investable and noninvestable ratings, and adding + or – in the same rating level represents the default strength. Table 3 shows the high-grade, intermediate-grade, low-grade, and lowest-grade credit rating indices. 4-level credit rating and 9-level risk rating were used as the experimental forecasting models. In this study, forecasting changes were observed after inputting DD as a variable into various credit rating models.

**Table 3 Credit rating definition indexes**

| 4- level credit rating | 9- Level risk rating | Definition | S&P |
|---|---|---|---|

| | | | |
|---|---|---|---|
| High grade | Prime | The debtor is highly capable of meeting financial commitments, but somewhat susceptible to adverse economic conditions and changes in circumstances. | AAA |
| | High grade | | AA+<br>AA<br>AA- |
| Intermediate grade | Upper medium grade | The debtor is adequately capable of meeting financial commitments, but highly susceptible to adverse economic conditions. | A+<br>A<br>A- |
| | Lower medium grade | | BBB<br>BBB<br>BBB- |
| Low grade | Noninvestment grade speculative | The debtor exhibits low vulnerability to adverse business, financial, and economic conditions, but is currently capable of meeting financial commitments. | BB+<br>BB<br>BB- |
| | Highly speculative | | B+<br>B<br>B- |
| | Substantial risks | | CCC<br>CCC<br>CCC- |
| Lowest grade | Extremely speculative | A bankruptcy petition has been filed or similar action has been taken; payment default on financial commitments might occur. | CC |
| | Default | | C |

### 3.3 Distance to Default Based on the KMV Model

This study involved using the KMV model (1989) to calculate the assets and volatility of enterprises. Maximizing debtor’s benefits were used to calculate the default point of the grace period after a payment due date and to analyze the default rate after the grace period, adopting the DD as a variable acquired from the KMV model. The two principal steps involved in calculating the DD are as follows:

$$V_E = V_A N(d_1) - De^{-r_f t} N(d_2) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

$$d_1 = -\frac{\ln\left(\frac{V_A}{D}\right) + \left(r_f + \frac{\sigma_A^2}{2}\right)t}{\sigma_A \sqrt{t}}$$

$$d_2 = d_1 - \sigma_A \sqrt{t}$$

$$\sigma_E = \frac{V_A \sigma_A N(d_1)}{V_E} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$\mathrm{DD} = \left[\frac{\ln\frac{V_A}{F} + \left(r - \frac{\sigma_A^2}{2}\right)t}{\sigma_A\sqrt{t}}\right] \quad (3)$$

where $V_A$ is the firm asset value,

$V_E$ is the market value of firm equity,

$r_f$ is the risk-free interest rate,

$F$ is the book value of firm liabilities,

$T$ is the time of the debt measurement,

$б_A$ is the asset value volatility of a firm,

$б_E$ is the equity and asset volatility of a firm,

$D$ is firm debt, and

$DD$ is the distance to default.

### 3.4 Variable and Feature Selection

In this study, all financial variables and the KMV model were based on data from the CompuStat North America dataset. North American financial and credit rating data from 2003 to 2012 were used as samples in this study. The North American enterprises that used the credit rating services of international credit rating agencies were explored, and companies that exhibited incomplete information or condition mismatching were eliminated. The dataset comprised 6,750 records, and 400 financial variables were selected from CompuStat for use in this study. These 400 variables of financial data from the annual CompuStat North American Fundamentals were processed during the first step in feature selection based on a four-level credit rating and nine-level risk rating; the feature selection was performed using the R language in an RF model. The mtryFactor and nTree parameters are critical in feature selection; accordingly, mtryFactor (= $\sqrt{400}$) represented that 20 variables were selected for comparison every time, nTree represented the number of DTs in the RF model. When nTree is well developed, it can be used to observe the level of out-of-bag error and difference in the number of DTs. When the errors stabilize, the nTree value can be determined. Figures 2 and 3 show the relationship between the out-of-bag error in the four-level credit rating and nine-level risk rating; this relationship represents the error rate of each level in the RF model.

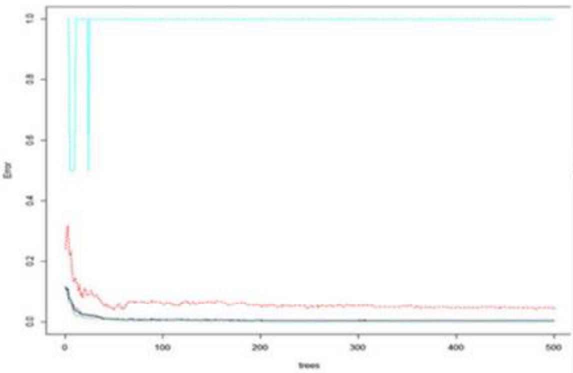

**Figure 2 out-of-bag error ( 4-level credit rating)**

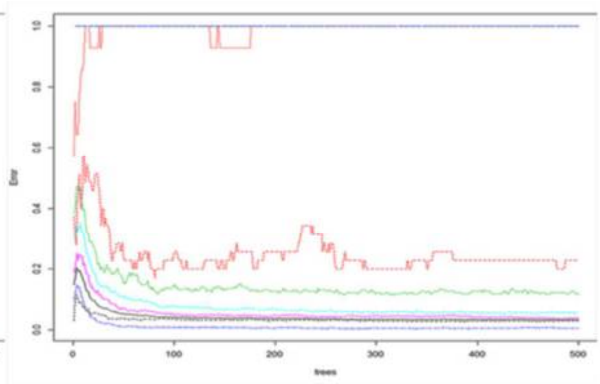

**Figure 3 out-of-bag error (9-level risk rating)**

The tenth (Figure 4) and fourth (Figure 5) DDs in both the four-level credit rating and nine-level risk rating based on the KMV model were used as the critical variables. The DDs were the necessary variables used in the variable evaluation. The validity of the DD is demonstrated in Figures 4 and 5, which prove the effectiveness of using the DD to predict credit ratings. The developed models were compared, and variables were added or subtracted to determine which variables influenced the accuracy rate. Figures 4 and 5 show the synthesis ranking of the selected variables.

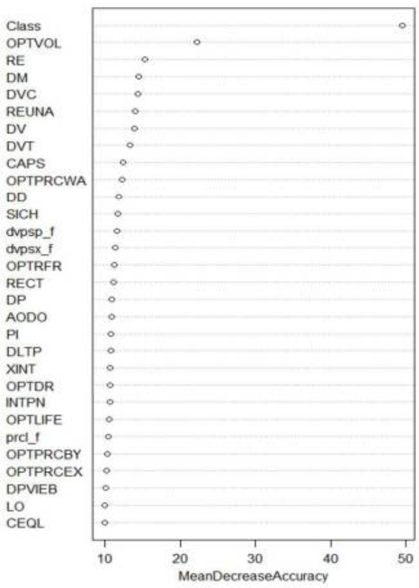


Figure 4 Variables ranking of 4-level credit rating

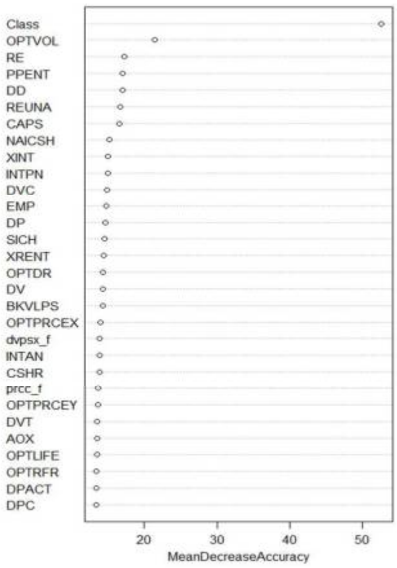


Figure 5 Variables ranking of 9-level risk rating

Regarding the feature selection performed using the RF model, 17 variables were selected from the two rating models (Table 4): the four-level credit rating and nine-level risk rating models. Most variables were related to the stock market, but others were related to charges, assets, and debts, indicating that stock profits and stock value play critical roles in credit rating and that investors are sensitive to stock price volatility in the capital market. In addition, the related depreciation expenses, assets, and debts are key indicators that affect credit ratings.

**Table 4 17 Variables chosen from the 2 rating models**

| Item No | Variable Code | Variable Name |
|---|---|---|
| 1 | SICH | Standard Industrial Classification |
| 2 | DD | Distance to Default |
| 3 | DP | Depreciation and Amortization |
| 4 | INTPN | Interest Paid Net |
| 5 | XINT | Interest and Related Expense |
| 6 | CAPS | Capital Surplus/Share Premium Reserve |
| 7 | DV | Cash Dividends (Cash Flow) |
| 8 | DVC | Dividends Common/Ordinary |
| 9 | dvpsx_f | Dividends per Share |
| 10 | DVT | Dividends - Total |
| 11 | OPTDR | Dividend Rate - Assumption (%) |
| 12 | OPTLIFE | Life of Options - Assumption |
| 13 | OPTPRCEX | Options Exercised - Price |
| 14 | OPTRFR | Risk Free Rate - Assumption (%) |
| 15 | OPTVOL | Volatility - Assumption (%) |
| 16 | RE | Retained Earnings |
| 17 | REUNA | Retained Earnings - Unadjusted |

## 4. RESULTS

### 4.1 Experimental Results

In this study, the RF model was employed to perform feature selection, and selected variables were input into the RF, DT, NN, and SVM classifiers to conduct data analyses.

Tables 5–8 show the results of the classifications based on each parameter set. In addition, a confusion matrix was employed as the evaluation standard in the classification. When the predicted values match the actual values, all cases from the models are classified into classification indexes (Bradley, 1997). Accuracy was defined as the number of correctly classified items divided by the total number of data items. The area under the receiver operating characteristic curve (ROC) was used to evaluate the level of classifier performance. A specific value ratio (AUC) approaching 1 indicated that the model exhibited an optimal forecasting performance level.

#### 4.1.2 Random forest

Three parameters were set in this study: numTrees, numFeatures and K (n folds cross-validation). The result of the matching analysis yielded 18 different training models, which were analyzed using the RandomForest classifier. The results of trainings were divided into 4-level and 9 classification predictions. Table 5 shows the classification results. The RF model was used to perform the feature selection process, in which variables were selected for classification. The classification result was tested using cross-validation to construct an optimal model. The results indicated that the 4-level credit rating was most accurate after undergoing feature selection. The highest accuracy rate was 95.5%, and the 9-level risk-rating yielded an 87.8% rate, suggesting that the RF model yields excellent forecasts. The results shown in Table 5 indicate that increasing the number of folds in cross-validation causes the better level of forecasting performance. The optimal parameter settings were ntree = 100 and numFeatures = 5 in the 10-fold cross-validation of the 4-level classification; by contrast, the optimal parameter settings were ntree = 250 and numFeatures = 5 in the 10-fold cross-validation of the 9-level classification.

**Table 5 Random Forest**

| Parameter 1 | Parameter 2 | Parameter 3 | 4-level classification | | | 9-level classification | | |
|---|---|---|---|---|---|---|---|---|
| | | | OOB | ROC Area | Accuracy | OOB | ROC Area | Accuracy |
| K=3 | ntree=100 | numFeatures =5 | 0.045 | 0.989 | 0.948 | 0.158 | 0.964 | 0.823 |
| | ntree=100 | numFeatures =6 | 0.047 | 0.989 | 0.944 | 0.164 | 0.963 | 0.822 |
| | ntree=250 | numFeatures =5 | 0.045 | 0.990 | 0.947 | 0.151 | 0.965 | 0.829 |
| | ntree=250 | numFeatures =6 | 0.045 | 0.990 | 0.946 | 0.152 | 0.965 | 0.825 |
| | ntree=500 | numFeatures =5 | 0.043 | 0.990 | 0.947 | 0.147 | 0.966 | 0.830 |
| | ntree=500 | numFeatures =6 | 0.043 | 0.990 | 0.946 | 0.153 | 0.966 | 0.827 |
| K=5 | ntree=100 | numFeatures =5 | 0.045 | 0.99 | 0.952 | 0.158 | 0.97 | 0.838 |
| | ntree=100 | numFeatures =6 | 0.047 | 0.989 | 0.951 | 0.164 | 0.969 | 0.836 |
| | ntree=250 | numFeatures =5 | 0.045 | 0.990 | 0.952 | 0.151 | 0.971 | 0.839 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | ntree=250 | numFeatures =6 | 0.045 | 0.990 | 0.952 | 0.152 | 0.97 | 0.839 |
| | ntree=500 | numFeatures =5 | 0.043 | 0.990 | 0.954 | 0.147 | 0.971 | 0.839 |
| | ntree=500 | numFeatures =6 | 0.043 | 0.986 | 0.952 | 0.153 | 0.971 | 0.838 |
| K=10 | ntree=100 | numFeatures =5 | 0.045 | 0.99 | 0.955 | 0.119 | 0.977 | 0.876 |
| | ntree=100 | numFeatures =6 | 0.047 | 0.991 | 0.954 | 0.123 | 0.976 | 0.874 |
| | ntree=250 | numFeatures =5 | 0.045 | 0.992 | 0.954 | 0.116 | 0.978 | 0.877 |
| | ntree=250 | numFeatures =6 | 0.045 | 0.992 | 0.953 | 0.116 | 0.977 | 0.875 |
| | ntree=500 | numFeatures =5 | 0.043 | 0.992 | 0.954 | 0.116 | 0.99 | 0.877 |
| | ntree=500 | numFeatures =6 | 0.043 | 0.992 | 0.953 | 0.119 | 0.977 | 0.876 |

### 4.1.3 Decision tree

Establishing the minNumObj (the minimun number of instance in a leaf), including K folds cross-validation, yielded 12 different trainings which were analyzed using the J48 classifier. Two types of classifications (i.e., the 4-level credit rating and 9-level risk rating) were used to produce training results. Table 6 shows the classification results.

The DT was used in the classification prediction. The results were tested using 3 different folds in cross-validation to construct the model. The classification results indicated that the four-level credit rating was most accurate (89.4%) after undergoing feature selection. The highest level of accuracy yielded by the nine-level risk rating was 73.8%.

The optimal parameter setting was minNumObj = 2 in the 10-fold cross-validation of the 4-level classification, and the optimal parameter setting was minNumObj = 2 in the 10-fold cross-validation of the 9-level classification.

**Table 6 Decision Tree**

| parameter 1 | parameter 2 | 4-level classification | | 9-level classification | |
|---|---|---|---|---|---|
| | | ROC Area | Accuracy | ROC Area | Accuracy |
| **K=3** | minNumObj=2 | 0.897 | 0.884 | 0.831 | 0.705 |
| | minNumObj=5 | 0.914 | 0.885 | 0.859 | 0.700 |
| | minNumObj=10 | 0.934 | 0.876 | 0.882 | 0.679 |
| | minNumObj=20 | 0.934 | 0.871 | 0.887 | 0.680 |
| **K=5** | minNumObj=2 | 0.897 | 0.889 | 0.845 | 0.731 |
| | minNumObj=5 | 0.924 | 0.884 | 0.872 | 0.719 |
| | minNumObj=10 | 0.937 | 0.886 | 0.890 | 0.708 |
| | minNumObj=20 | 0.934 | 0.876 | 0.895 | 0.693 |

| | | | | | |
|---|---|---|---|---|---|
| **K=10** | minNumObj=2 | 0.903 | 0.894 | 0.853 | 0.738 |
| | minNumObj=5 | 0.931 | 0.893 | 0.88 | 0.726 |
| | minNumObj=10 | 0.940 | 0.888 | 0.892 | 0.708 |
| | minNumObj=20 | 0.938 | 0.878 | 0.895 | 0.694 |



### 4.1.4 Neural network

Setting hidden layers (H), learning rate (L), momentum (M), and training time (T) as variables, including K folds cross-validation, yielded 24 different trainings, which were analyzed using the MultilayerPerceptron classifier. The 24 trainings were divided into two models for classification: a 4-level credit rating and 9-level risk rating. Table 7 lists the classification results. The back propagation neural network was used to perform the classification, and the results were tested using 3 different folds in cross-validation to construct the model. The results indicated that the 4-level credit rating was most accurate (88.4%) after performing classification by using the NN. The highest rate of accuracy yielded by the 9-level risk rating was 66.2%. The optimal parameter settings were H = 10, L = 0.1, M = 0.2, and T = 500 in the three-fold cross-validation of the 4-level classification.

**Table 7 Neural network**

| Parameter 1 | Parameter 2 | Parameter 3 | Parameter 4 | Parameter 5 | 4-level classification | | 9-level classification | |
|---|---|---|---|---|---|---|---|---|
| | | | | | ROC | Accuracy | ROC Area | Accuracy |
| **K=3** | H=5 | L=0.1 | M=0.2 | T=500 | 0.943 | 0.871 | 0.866 | 0.635 |
| | H=5 | L=0.1 | M=0.4 | T=500 | 0.94 | 0.875 | 0.866 | 0.632 |
| | H=5 | L=0.3 | M=0.2 | T=500 | 0.94 | 0.877 | 0.87 | 0.636 |
| | H=5 | L=0.3 | M=0.4 | T =500 | 0.938 | 0.871 | 0.859 | 0.625 |
| | H=10 | L=0.1 | M=0.2 | T =500 | 0.942 | 0.873 | 0.877 | 0.645 |
| | H=10 | L=0.1 | M=0.4 | T =500 | 0.946 | 0.884 | 0.874 | 0.644 |
| | H=10 | L=0.3 | M=0.2 | T =500 | 0.944 | 0.882 | 0.864 | 0.631 |
| | H=10 | L=0.3 | M=0.4 | T =500 | 0.943 | 0.878 | 0.863 | 0.636 |
| **K=5** | H=5 | L=0.1 | M=0.2 | T =500 | 0.942 | 0.876 | 0.875 | 0.650 |
| | H=5 | L=0.1 | M=0.4 | T =500 | 0.941 | 0.871 | 0.874 | 0.645 |
| | H=5 | L=0.3 | M=0.2 | T =500 | 0.936 | 0.867 | 0.87 | 0.646 |
| | H=5 | L=0.3 | M=0.4 | T =500 | 0.934 | 0.874 | 0.869 | 0.637 |
| | H=10 | L=0.1 | M=0.2 | T =500 | 0.943 | 0.876 | 0.881 | 0.663 |
| | H=10 | L=0.1 | M=0.4 | T =500 | 0.941 | 0.872 | 0.879 | 0.655 |
| | H=10 | L=0.3 | M=0.2 | T =500 | 0.943 | 0.879 | 0.876 | 0.654 |
| | H=10 | L=0.3 | M=0.4 | T =500 | 0.938 | 0.872 | 0.872 | 0.653 |
| **K=10** | H=5 | L=0.1 | M=0.2 | T =500 | 0.941 | 0.872 | 0.879 | 0.651 |
| | H=5 | L=0.1 | M=0.4 | T =500 | 0.939 | 0.871 | 0.877 | 0.646 |
| | H=5 | L=0.3 | M=0.2 | T =500 | 0.936 | 0.868 | 0.868 | 0.637 |
| | H=5 | L=0.3 | M=0.4 | T =500 | 0.935 | 0.869 | 0.87 | 0.639 |

| Parameter 1 | Parameter 2 | Parameter 3 | Parameter 4 | Parameter 5 | 4-level classification | | 9-level classification | |
|---|---|---|---|---|---|---|---|---|
| | | | | | ROC | Accuracy | ROC Area | Accuracy |
| | H=10 | L=0.1 | M=0.2 | T =500 | 0.945 | 0.882 | 0.884 | 0.662 |
| | H=10 | L=0.1 | M=0.4 | T =500 | 0.946 | 0.879 | 0.88 | 0.655 |
| | H=10 | L=0.3 | M=0.2 | T =500 | 0.943 | 0.878 | 0.877 | 0.653 |
| | H=10 | L=0.3 | M=0.4 | T =500 | 0.939 | 0.875 | 0.87 | 0.646 |

**H:** hidden layers, **L:**learning rate, **M:**momentum, **T:**training time K: n folds cross-validation

**4.1.5 Support vector machine.**

Using PolyKernel as the default setting variable, including K folds cross-validation, yielded three different trainings, which were analyzed using the SMO classifier. The 3 trainings were divided into a 4-level credit rating and nine-level risk rating for classification. Table 8 shows the classification results. The SVM was used to conduct the classification. The results were tested using 3 different folds in cross-validation to construct the model. The results indicated that the four-level credit rating model was most accurate (88.1%) after performing a classification by using the SVM. The highest accuracy rate yielded by the nine-level risk rating model was 67.3%. The optimal parameter setting was in the 10-fold cross-validation of the four-level and nine-level classifiers.

**Table 8 Support vector machine**

| parameter 1 | parameter 2 | 4-level classification | | 9-level classification | |
|---|---|---|---|---|---|
| | kernel | ROC Area | Accuracy | ROC Area | Accuracy |
| **K=3** | PolyKernel | 0.869 | 0.867 | 0.79 | 0.659 |
| **K=5** | PolyKernel | 0.870 | 0.877 | 0.793 | 0.665 |
| **K=10** | PolyKernel | 0.871 | 0.881 | 0.794 | 0.673 |

Regarding the parameter setting in the four classifiers, the results indicated that the RF classifier was most accurate, followed by the DT and NN classifiers. The SVM exhibited the least favorable performance. Similarly, in the both 4-level classification and 9-level classification, the RF exhibited superior forecasting capacity. Table 9 lists the results, indicating that, regardless of the classifier, the accuracy of the variables obtained using feature selection was more favorable compared with that of the variables obtained without using feature selection; this validates the effectiveness of feature selection, implying that the RF model is the most suitable classifier.

**Table 9 Classifiers Prediction Rate Based on Feature Selection**

| **4 level credit rating** | | **Random Forest** | **Decision Tree** | **Neural network** | **Support vector machine** |
|---|---|---|---|---|---|
| **Feature Selection Using Random Forest** | parameter | k=10<br>nT=100,nF =5 | k=10<br>mNO=2 | k=10<br>H=10,L=0.1<br>M=0.4,T =500 | k=10<br>kernel=PolyKernel |
| YES | ROC Area | 0.990 | 0.903 | 0.946 | 0.871 |
| | Accuracy | 0.955 | 0.894 | 0.884 | 0.881 |
| NO | ROC Area | 0.986 | 0.897 | 0.935 | 0.893 |
| | Accuracy | 0.927 | 0.893 | 0.873 | 0.857 |
| **9 Level risk rating** | | **Random Forest** | **Decision Tree** | **Neural network** | **Support vector machine** |
| **Feature Selection Using Random Forest** | parameter | k=10<br>nT=250,nF =5 | k=10<br>mNO=2 | k=10<br>H=10,L=0.1<br>M=0.2,T =500 | k=10<br>kernel=PolyKernel |
| YES | ROC Area | 0.978 | 0.853 | 0.884 | 0.793 |
| | Accuracy | 0.877 | 0.738 | 0.662 | 0.673 |
| NO | ROC Area | 0.971 | 0.829 | 0.870 | 0.831 |
| | Accuracy | 0.828 | 0.707 | 0.648 | 0.625 |

## 5. CONCLUSION

The research data were obtained from CompuStat North America. The financial data options were limited by the completeness of the data. In the data-mining field, the RF is a novel prediction model. The RF model can be used in classification and variable feature selection to simplify complex financial variable structures. This research involved evaluating credit rating models by using data-mining techniques. The RF model exhibited a more favorable level of performance than did current models, namely the DT, SVM, and NN models. The RF model attained an accuracy rate of 95.5% in the 4-level classification and 87.7% in the 9-level classification, representing a superior level of performance compared with that achieved by the other three models. The results indicated that the RF model can be used to perform feature selection and efficiently filter numerous financial variables. In addition, the research findings indicated the most effective financial variables were dividends common/ordinary, cash dividends, volatility assumption, and risk-free rate assumption in this study. Future studies can employ financial data from various areas and compare how these areas influence the forecasting model, adding feature selection variables. Additional explanatory variables and effective models for use in credit rating prediction warrant further discussion.

## REFERENCES


Altman, E. &Saunders, A. (1997) "Credit risk measurement: Developments over the last 20 years", *Journal of Banking and Finance*,Vol. 21,No.11-12, pp. 1721-1742.

Angelini, E., di Tollo, G.& Roli, A. (2008) "A neural network approach for credit risk evaluation, *The Quarterly Review of Economics and Finance* ,Vol. 48, No.4, pp. 733-755.

Bailey, M., Betts, J., Cowie, P., Dekker, B., Denby, S., Hoyland, C., Kindred, D., Lucas, A., Neves, E., Pearce, W., Poole, D. & Trupp, S. (2004) *Credit Scoring: the Principles and Practicalities*, White Box Publishing.

Beaver, W. H. (1966) "Financial Ratios As Predictors of Failure", *Journal of Accounting Research* Vol. 4, pp.71-111.

Bharath, S. T. & Shumway, T. (2008) "Forecasting Default with the Merton Distance to Default Model", *Review of Financial Studies* ,Vol. 21, No. 3, pp.1339-1369.

Black, F. & Myron, S. (1973) "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy* ,Vol.81, No. 3, pp. 637-654.

Bradley, A. P. (1997) "The use of the area under the ROC curve in the evaluation of machine learning algorithms", *Pattern Recognition*, Vol. 30, No. 7, pp. 1145-1159.

Breiman, L. (2001) "Random Forests", *Mach. Learn*,Vol. 45, No. 1, pp. 5-32.

Campbell, J. E., Lobell, D. B., Genova, R. C & Field, C. B. (2008) "The global potential of bioenergy on abandoned agriculture lands", *Environ Sci Technol*, Vol. 42, No. 15, pp. 5791-5794.

Chen, W. H. & Shih J. Y.,(2006) "A study of Taiwan's issuer credit rating systems using support vector machines", *Expert Systems with Applications*, Vol. 30, No. 3, pp. 427-435.

Chen, Y.S. & Cheng, C.H. (2013) "Hybrid models based on rough set classifiers for setting credit rating decision rules in the global banking industry", *Knowledge-Based Systems* , Vol. 39, pp. 224-239.

Díaz-Uriarte, R. & De Andres, S. A. (2006) "Gene selection and classification of microarray data using random forest", *BMC bioinformatics*, Vol. 7, No. 1, pp.3.

Du, Y. & Suo, W. (2007) "Assessing credit quality from the equity market: can a structural approach forecast credit ratings?", *Canadian Journal of Administrative Sciences / Revue Canadienne des Sciences de l'Administration* ,Vol. 24, No. 3, pp. 212-228.

Dutta, S. & Shekhar, S. (1988) *Bond rating: a nonconservative application of neural networks*, 1988., IEEE International Conference on Neural Networks,.

Gupton, G. M., Morgan, J. P.,Co, C., Finger, C.& Bhatia, M. (1997) *CreditMetrics, Technical Document: The Benchmark for Understanding Credit Risk*, J.P. Morgan.

Kliger, D. & Sarig, O. (2000) “The information value of bond ratings”, *Journal of Finance* ,Vol. 55, No. 6, pp. 2879-2902.

Lopez, J. A. (2004) “The empirical relationship between average asset correlation, firm probability of default, and asset size”, *Journal of Financial Intermediation*, Vol. 13,No. 2, pp. 265-283.

Maher, J. J. & Sen, T. K. (1997) “Predicting Bond Ratings Using Neural Networks: A Comparison with Logistic Regression”, *International Journal of Intelligent Systems in Accounting, Finance & Management* ,Vol. 6, No. 1, pp. 59-72.

Mays, E. (2001). *Handbook of Credit Scoring*, Stylus Pub Llc.

Mays, E. & Mays, F. E. (1998) *Credit Risk Modeling: Design and Application*, Glenlake.

Merton, R. C. (1974) “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates”, *The Journal of Finance* ,Vol. 29, No.2, pp. 449-470.

Mori, H. & Umezawa, Y. (2007) *Credit risk evaluation in power market with random forest*, , 2007. ISIC. IEEE International Conference on Systems, Man and Cybernetics.

Odom, M. D. & Sharda, R. (1990) *A neural network model for bankruptcy prediction*, , 1990., IJCNN International Joint Conference on Neural Networks

Shin, K. S. & Han, I. (2001) “A case-based approach using inductive indexing for corporate bond rating”, *Decision Support Systems* ,Vol. 32, No. 1, pp. 41-52.

Strobl, C., Boulesteix, A. L., Kneib, T., Augustin, T.& Zeileis, A. (2008) “Conditional variable importance for random forests”, *BMC bioinformatics* ,Vol. 9, pp. 307.

Härdle, W., Prastyo, D & Hafner, C. (2012). “ Support Vector Machines with Evolutionary Feature Selection for Default Prediction”, Sonderforschungsbereich 649, Humboldt University, Berlin, Germany.

Yeh, C.-C., Chi, D.-J. & Hsu, M.-F. (2010) “A hybrid approach of DEA, rough set and support vector machines for business failure prediction”, *Expert Systems with Applications* ,Vol. 37, No. 2, pp. 1535-1541.

Yeh, C. C., Lin, F. Y. & Hsu, C. Y. (2012) “A hybrid KMV model, random forests and rough set theory approach for credit rating”, *Knowledge-Based Systems*, Vol. 33, pp. 166-172.

Zhou, J.& Bai, T. (2008) “Credit Risk Assessment Using Rough Set Theory and GA-Based SVM”, *Proceedings of the 2008 The 3rd International Conference on Grid and Pervasive Computing - Workshops*, IEEE Computer Society, pp. 320-325.