Credit Risk Estimation Model Development Process: Main Steps and Model Improvement

Ricardas Mileris, Vytautas Boguslauskas

Kaunas University of Technology
K. Donelaičio str. 73, LT-44309, Kaunas, Lithuania
e-mail: ricardas.mileris@ktu.lt, vytautas.boguslauskas@ktu.lt

crossref http://dx.doi.org/10.5755/j01.ee.22.2.309

The attribution of credit ratings for clients is a very important issue in the banking sector. Banks must evaluate credit risk of credit applicants by using standardized (external rating institutions) or internal ratings-based (IRB) methods. Banks which decided to use IRB method attempt to develop precise internal credit rating models for the evaluation of credit worthiness of their borrowers. The internal rating method for the estimation of default probability requires to collect the default information from the historical data in banks. The major studies about default determining factors are based on classification methods (Zhou, Xie, Yuan, 2008). A classification model considers the default measurement as the pattern recognition where all borrowers are divided into non-default and default groups based on their financial and non financial data. Banks attempt to construct an evaluation model that can be used to discriminate new sample.

This research focuses on a credit rating model development which could attribute credit ratings for Lithuanian companies. The steps of a model’s development and improvement process are described in this paper.

The model’s development begins with the selection of initial variables (financial ratios) characterizing default and non-default companies. 20 financial ratios of 5 years were calculated according to annual financial reports. Then statistical and artificial intelligence methods were selected for the classification of companies into two groups: default and non-default. A discriminant analysis, logistic regression and artificial neural networks (multilayer perceptron) were applied for this purpose. Often statistical methods are not able to operate with a large amount of data, so the analysis of variance, Kolmogorov-Smirnov test and factor analysis were applied for data reduction. Artificial neural networks often are able to analyze a large amount of data so variable selection was accomplished by the network itself calculating ranks of importance for every initial variable. There were constructed 15 classification models and their classification accuracy was measured by calculating correct classification rates. The most accurate was a logistic regression model analyzing data of 3 years (97% of correctly classified companies). Then the sample of companies was supplemented with new data and changes in classification accuracy were estimated. The significant decrease of classification accuracy conditioned the need of model update. For this reason the logistic regression coefficients were recalculated. In order to classify non-default companies into 7 classes: profitability, liquidity, financial structure and individual possibility of default estimated by a logistic regression model were determined as rating criterions. Then the rating scale was constructed and credit ratings were attributed for companies in the sample. The calculated probabilities of default indicated that some lower ratings have lower probabilities of default. These imperfections were corrected by the modification of a rating scale. The research has shown that the developed model is a valid tool for the estimation of credit risk.

Keywords: bank, credit ratings, credit risk, classification methods.

Introduction

Banking activity is exposed to various risks. Understanding and quantifying these risks is crucial for bank management as well as for stability of the whole economy (Siecza, Holyst, 2009). Banks tend to lend to firms with high credit quality and not to lend to low credit quality firms. So the most important factor in determining lending practices is credit risk (Daniels, Ramirez, 2008). According to Dufl & Einig (2009), research considering credit risk has been one of the most active areas of recent financial research, with significant efforts deployed to analyse the meaning, role, and influence of credit ratings.

The development of the credit rating prediction models has attracted lots of research interests in academic and business community. Many researchers have attempted to construct automatic classification systems using methods from data mining, such as statistical and artificial intelligence techniques (Huang, 2009). These techniques include discriminant analysis (Min, Lee, 2008), logistic regression (Psillaki, Tsolas, Margaritis, 2010), k-nearest neighbour (Twala, 2010), decision trees (Paleologo, Elisseeff, Antonini, 2010), artificial neural networks (Angelini, Tollo, Roli, 2008; Yu, Wang, Lai, 2008), support vector machine (Danenas, Garsva, 2009; Lee, 2007), nearest subspace method (Zhou, Jiang, Shi, 2010), hybrid models (Lin, 2009), etc.

Mostly the scientific publications about estimation of credit risk present models which classify companies into 2 groups: default or non-default. This research is intended for the development of credit risk estimation model which could classify reliable companies into seven classes and not reliable companies into one class.

The object of this research is credit risk estimation models.

The aim of this research is to develop credit risk estimation model and to describe the development process.
The methods of this research:
1. Analysis of scientific publications.
2. Development of a credit risk estimation model.
   The developed model will allow the bank to estimate the credit risk of Lithuanian companies.

Internal credit rating models in banks

The activity of business companies is directly or indirectly influenced by various internal and external factors (Boguslauskas, Adlyte, 2010). Unfavourable business environment, unexpected and unfavourable events, the risky decisions of enterprise managers may make any enterprise insolvent and lead it towards bankruptcy (Purvinis, Sukys, Virbickaite, 2005). Difficult business environment stimulates to look for new ways to assess the changing situation as well as to implement and manage new means for business continuity (Valackiene, 2010). These negative factors and ability to manage them also influence the credit risk of companies which must be measured by banks properly.

With the rapid growth in credit industry and the management of large loan portfolios, internal credit rating models have been extensively used for the credit admission evaluation. In general, the bank’s internal measures of credit risk are based on assessments of borrower and transaction risk. Mostly banks base their rating methodology on borrower's default risk and typically assign a borrower to a credit rating grade (Butera, Faff, 2006). The credit rating models are developed to classify loan applicants as either a reliable group (accepted) or not reliable group (rejected) with their related characteristics based on the data of the previous accepted and rejected applicants (Tsai, Chen, 2010).

The guidelines for measuring credit risk are suggested in the Basel II Capital Accord (Weißbach, Tschiersch, Lawrenz, 2009). With the implementation of the Basel II Capital Accord in 2006, banks are now also allowed to use internal data and rating models for the purpose of estimating own default probabilities (PD) and calculating regulatory capital (Ryser, Denzler, 2009). According to the Basel Committee on Banking Supervision banks are required to measure the one year default probability. So usually statistical credit risk estimation models try to predict the probability that a loan applicant or existing borrower will default in one year (Fantazzini, Figini, 2009). In addition, banks should also provide the loss given default (LGD) – a measure of how much per unit of exposure bank will lose if the client defaults (Butera, Faff, 2006). Credit rating models allow to rate clients credibility and to evaluate expected losses of bank (Sieczka, Holyst, 2009).

One of the central ideas of IRB method that banks are free to choose their own method when:
1. A meaningful differentiation of risk between classes is achieved.
2. Plausible, intuitive, and current input data are used.
3. All relevant information is taken into account (Ryser, Denzler, 2009).

Commercial banks can use statistical tools to discriminate between reliable and not reliable borrowers and to measure the credit quality according to internal rating scales (Ryser, Denzler, 2009). There are several well-known statistical techniques for constructing credit rating predictions. These techniques include logistic regression, discriminant analysis, linear probit model, etc. (Hwang, Chung, Chu, 2010). Also methods of artificial intelligence are widely used for classification of companies. In addition, the hybridization approach has been an active research area to improve the classification (prediction) performance over single learning approaches. In general, it is based on combining two different machine learning techniques. Therefore, to develop a hybrid learning credit model, there are three different ways to combine the two machine learning techniques:
1. Combining two classification techniques.
2. Combining two clustering techniques.
3. One clustering technique combined with one classification technique (Tsai, Chen, 2010).

These statistical and machine learning techniques analyze the particular data set about credit applicants. So in credit risk estimation model development the successful selection of informative variables is very important. According to Arslan & Karan (2009), different factors affect credit risks of firms. Jiao, Syau & Lee (2007) maintained that the data can be classified into three categories:
1. Financial conditions: liquidity, financial structure, profitability and efficiency ratios.
2. Management measures: administrator’s management experiences, stockholders structure type, conditions of capital increment during the last years, etc.
3. Characteristics and perspectives of products and competitions: equipment and technologies, product marketability, collateral, economic conditions of the industry in the next year.

Thus, the total credit rating can be obtained by summarizing these three categories. It should be noted that only the first category (financial conditions) can be expressed numerically. Most of the quantities in the other two categories cannot be expressed numerically easily. To overcome this difficulty, fuzzy numbers can be used for linguistic expressions (Jiao, Syau, Lee, 2007).

The interaction between different risk factors and credit ratings was analyzed in many publications. The importance of financial factors is widely accepted because its impact is measurable. The relevance of non-financial factors is mainly considered as not decisive (Grüntert, Norden, Weber, 2005). So companies should seek to present a true and fair view of their financial performance and results because banks need informative and truthful accounting data for making right decisions (Rudzioniene, 2006).

The results of a credit risk estimation are credit ratings. Usually credit ratings range from AAA to D. Basing on the rating scale of Standard & Poor’s, it is possible to group ratings into three categories:
- AAA, AA, A as category 1.
- BBB as category 2.
- Below BBB as category 3.

Firms in the {AAA, AA, A} category mean that they have demonstrated strong capacity to meet their financial obligations. Firms receiving BBB rating mean that they
have adequate capacity to meet their financial commitments. However, firms receiving ratings below BBB mean that they are regarded as having speculative characteristics (Hwang, Chung, Chu, 2010).

The internal rating models have many benefits for banks. These benefits include reducing the cost of credit analysis, enabling faster decisions, ensuring credit collections, and diminishing possible risks. Even a slight improvement in credit rating accuracy may reduce large credit risks and translate into significant future savings (Tsai, Chen, 2010). In addition Ince & Aktan (2009) affirm that credit rating models, used to model the potential risk of loan applications, may be an effective substitute for the use of judgment among inexperienced loan officers, thus helping to control bad debt losses.

When the credit risk estimation model is developed, it is important to evaluate its quality. The low-quality rating models have two important negative effects on a bank’s financial stability. First, return on the portfolio will be lower than expected. If a borrower receives too favorable rating, he will most likely accept the loan offer, whereas rejection is more likely in the case of a rating error in the opposite direction, i.e., the bank will earn money from adverse selection. Second, due to the adverse selection effect, the calculated regulatory capital held by the bank is too low. As more borrowers with too favorable ratings are in the portfolio, the resulting regulatory capital based on these ratings may be inadequate (Hornik, Jankowitsch, Lingo, Pichler, Winkler, 2010).

**Credit risk estimation model development**

The model for rating of Lithuanian companies was developed in order to measure their credit risk. The main steps of a model’s development process are described below.

1. **Variable selection.** For typical classification problems, values for a set of independent variables are given in a set of (training) examples, upon which a model is developed to categorize future observations into appropriate classes (Piramuthu, 2006). In this case the analyst must have the set of variables which allow to estimate the credit risk of companies successfully. Developing the credit risk estimation model, the set of initial variables consisted of 20 financial ratios of 5 years. So there were 100 independent variables \( X_i \) and one dependent variable \( Y \) (default or non-default).

2. **Data analysis methods selection.** Discriminant analysis (DA), logistic regression (LR) and artificial neural networks (ANN) methods were applied for classification of companies into 2 groups: default and non-default.

   Discriminant analysis is a commonly used technique to classify a set of observations into predefined classes in various fields (Huang, Kao, Wang, 2007). It is a classification method that can predict the group membership of a newly sampled observation. In DA, a group of observations whose memberships are already identified are used for the estimation of weights in a discriminant function. A new sample is classified into one of several groups by DA results (Sueyoshi, 2006). The goal of linear DA is to find a linear transformation that maximizes class separability in the reduced dimensional space (Park, Park, 2008). The mathematical model is:

\[
    z = \alpha + \sum_{i=1}^{n} \beta_i x_i 
\]

where \( \alpha \) is an intercept, \( x_i \) is the value of independent variable, \( \beta_i \) is the coefficient of the variable \( x_i \).

Two discriminant functions were constructed for default (\( z_0 \)) and non-default (\( z_0 \)) companies in every developed DA model. The company was numbered into group which \( z \) gets higher value.

Logistic regression models allow to estimate the individual possibility of default for every company in the range of \([0; 1]\). In a logistic regression model the possibility of individual to default is expressed as follows:

\[
    P_n = \frac{\exp(\alpha + \sum_{i=1}^{n} \beta_i x_i)}{1 + \exp(\alpha + \sum_{i=1}^{n} \beta_i x_i)}
\]

where \( \alpha \) is an intercept, \( x_i \) is the value of independent variable, \( \beta_i \) is the coefficient of the variable \( x_i \) (Dong, Lai, Yen, 2010). The classification threshold of \( P_n \) in developed LR models was set to 0.5.

A multilayer perceptron was used for the classification of companies as one of the possible types of ANN (Figure 1).

**Figure 1. Multilayer perceptron**

In this network, the information flows forward to the output continuously without any feedback (Boguslauskas, Mileris, 2009). The initial variables are fed into input nodes, while the output provides the forecast for the future value. Hidden nodes with appropriate nonlinear transfer functions are used to process the information received by the input nodes. The model can be written as:

\[
    y_i = \alpha_o + \sum_{j=1}^{m} \alpha_j f \left( \sum_{l=1}^{l} \beta_{oj} x_l + \beta_{ij} \right) + \epsilon_i
\]

where \( n \) is the number of input nodes, \( m \) is the number of hidden nodes, \( f \) is a sigmoid transfer function, \( \{ \alpha_j, j = 0, 1, \ldots, m \} \) is a vector of weights from the hidden to output.
nodes, \( \{ \beta_{ij}, i = 1, 2, ..., n; j = 0, 1, ..., m \} \) are weights from the input to hidden nodes, \( \alpha_0 \) and \( \beta_0 \) are weights of connections leading from the bias which have values always equal to 1 (Azadeh, Ghaderi, Sohrabkhani, 2007).

3. Data reduction. Not all initial variables are useful for the analysis, so it is important to find informative parameters for credit risk estimation model development in the set of initial variables. Also the statistical methods often are not able to analyze high amount of data, so the need of data reduction arises. Four methods for data reduction were applied: the analysis of variance (ANOVA), Kolmogorov-Smirnov test (K-S), factor analysis (FA) and ranks of importance in ANN (ROI).

Figure 2 illustrates the process of data reduction for classification models. The 1st column consists of the period of data (from the last year to 5 years ago) and number of initial variables. The 2nd and 3rd columns show methods applied for data reduction. In the 4th column there is a number of variables used for classification of companies. The 5th column consists of a code of classification models (method and period of data analyzed).

**Figure 2.** Data reduction for classification models

The purpose of ANOVA is to test for significant differences between means of independent variables in groups of failed and not failed companies. If the means did not differ significantly the variable was rejected from further analysis.

The Kolmogorov-Smirnov one sample test for normality is based on the maximum difference between the sample cumulative distribution and the hypothesized cumulative distribution (Mileris, Boguslauskas, 2010). This test allowed to verify if the variable has the normal distribution. If the statistics D of K-S test is significant, then the hypothesis that the respective distribution is normal should be rejected. These variables were not included into the credit risk analysis.

Factor analysis is a statistical method that is based on the correlation analysis of variables. The purpose is to reduce multiple variables to a lesser number of factors (Ocal, Oral, Erdis, Vural, 2007). The reduction of observed initial variables to less factors allows to enhance interpretability and detect hidden structures in the data (Treiblmaier, Filzmoser, 2010).

These methods were applied for data reduction before DA and LR models were developed. Artificial neural networks are able to operate with a large amount of data, so all initial variables were fed to them. Ranks of importance were attributed for every variable by ANN itself. The variables that ROI > 0 were included into further analysis.

4. Estimation of classification accuracy. Overall 15 models were developed for the classification of companies into 2 groups. The correct classification rates (CCR) were calculated for the estimation of classification accuracy (Table 1). These rates indicate the proportion of correctly classified companies by a certain model.

The highest classification accuracy (97%) was reached by the logistic regression model which employed three-years financial data of companies. So this LR3 model was used for the estimation of credit risk.

5. Estimation of changes in classification accuracy. The classification model LR3 was developed using financial ratios of 2004-2006 years. The data of 100 Lithuanian companies was analyzed. After that the initial data set was complemented with financial ratios of 2006-2008 years of 100 new companies and the changes in classification accuracy were estimated. The test of model indicated that the correct classification rate decreased to 82.23% (Table 2). It is the significant decrease (\( \Delta \alpha = -14.77\% \)) so the need of model update arised.

6. Update of classification model. The regression coefficients were recalculated for the same 25 variables of LR3 model. The CCR of updated classification model increased to 93.43% (Table 2). The increase influenced on model update was 11.2% (\( \Delta \beta \)).

**Table 2**

<table>
<thead>
<tr>
<th>Changes of CCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR3 Test</td>
</tr>
<tr>
<td>97.00</td>
</tr>
</tbody>
</table>

Changes of CCR show the importance of model update when bank gets new data about clients. The update can significantly improve the classification model’s performance.

7. Determination of ratings criterions. The factor analysis has shown that 3 factors which explain 40.4% –
42.8% of total variance are: profitability, liquidity and financial structure. So from an initial set of variables 4 profitability, 2 liquidity and 1 financial structure ratios were selected which had the highest correlation coefficients \(|r|\) with individual possibility of default \((P_n)\).

### Table 3

| Group          | Financial ratio | \(r\)  | \(|r|\) |
|----------------|-----------------|--------|--------|
| Profitability: | NPM             | -0.27147 | 0.27147 |
|                | EBIT/TA         | -0.50944 | 0.50944 |
|                | NP/TA           | -0.48495 | 0.48495 |
|                | EBIT/S          | -0.31458 | 0.31458 |
| Liquidity:     | CR              | -0.15071 | 0.15071 |
|                | QR              | -0.12926 | 0.12926 |
| Financial structure: | DR         | 0.181803  | 0.181803 |

In Table 3 NPM is a net profit margin, EBIT/TA is earnings before interest and taxes to total assets, NP/TA is net profit to total assets, EBIT/S is earnings before interest and taxes to sales, CR is current ratio, QR is quick ratio, EBIT/TA is earnings before interest and taxes to sales, CR is current ratio, QR is quick ratio, DR is debt ratio. As a criterion of ratings the individual possibility of default also included into analysis. So the profitability determines 50%, liquidity – 25%, financial structure – 12.5%, individual possibility of default – 12.5% of credit rating.

8. Construction of rating scale. The rating model is based on classification of companies according to logistic regression analysis results and rating criterions (Figure 3).

The variables \(X_i\) were analyzed by the logistic regression (LR3) model. According to estimated individual possibility of default \((P_n)\), companies were classified into 2 classes: default and non-default. For companies classified by LR3 model as default the rating D1 was attributed. The Basel II Capital Accord requires that non-default companies must be classified into at least 7 classes. So according to the determined rating criterions ratings AAA – C were attributed for reliable companies. In addition, if company by LR3 model was classified as reliable but its financial ratios are bad, for such companies rating D2 was attributed.

In group of not default companies 5 statistical characteristics were calculated for 7 financial ratios: minimum value \((\text{Min})\), maximum value \((\text{Max})\), standard deviation \((\sigma)\), mean \((\mu)\) and median \((\text{Me})\). The exceptions \((E)\) of values were rejected from further analysis:

\[
E \notin (\mu - 3\sigma, \mu + 3\sigma) \quad (4)
\]

Then \(\text{Min}, \text{Max}\) and \(\text{Me}\) values in the sample were recalculated. In order to create rating scale the interval of values for each financial ratio was divided into 2 sections:

- \(\text{Min} - \text{Me}\) (from minimum to median).
- \(\text{Me} - \text{Max}\) (from median to maximum).

Every section was divided into 4 equal intervals and the margins of financial ratios in these intervals were calculated. According to the values of financial ratios and individual possibility of default \((P_n)\), scores were attributed for companies (Table 4).

### Table 4

<table>
<thead>
<tr>
<th>Score</th>
<th>NPM</th>
<th>EBIT/TA</th>
<th>...</th>
<th>(P_n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>&gt; 0.75</td>
<td>&gt; 0.38</td>
<td>...</td>
<td>(\leq 0.001)</td>
</tr>
<tr>
<td>6</td>
<td>(0.51; 0.75)</td>
<td>(0.28; 0.38)</td>
<td>...</td>
<td>(0.001; 0.003]</td>
</tr>
<tr>
<td>0</td>
<td>(\leq 0.01)</td>
<td>(-0.0772)</td>
<td>...</td>
<td>&gt; 0.541</td>
</tr>
</tbody>
</table>

The credit rating of a company depends on the sum of scores (Table 5).

### Table 5

<table>
<thead>
<tr>
<th>Rating</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>...</th>
<th>D2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total score</td>
<td>50-56</td>
<td>43-49</td>
<td>36-42</td>
<td>...</td>
<td>0-7</td>
</tr>
</tbody>
</table>

9. Calculation of probabilities of default. The probabilities of default \((PD)\) were calculated for every rating:

\[
PD = \frac{N_C}{I_C} \times 100\% \quad (5)
\]

where \(N_C\) – number of default companies in particular rating, \(I_C\) - number of companies in particular rating.

The probabilities of default illustrated in Figure 4.

Gruntet, Norden & Weber (2005) point out the necessity of a link between credit rating and probability of default. The credit rating models of high quality have the inverse relation between credit ratings and PD values: when the credit ratings decrease the PD values constantly increase. In developed model ratings A and BBB have higher PD than rating BB. Also it was impossible to calculate PD for rating AAA because no one company got this rating. So it is worth to modify the rating scale in order to improve the model’s performance.
10. Modification of rating scale. The intervals of total scores were changed (Table 5) and all imperfections were corrected (Figure 5).

![Figure 5. PD (%) and credit ratings of modified scale](image)

Companies with ratings D1 and D2 can be considered as one class of not reliable clients and bank should not credit them. These ratings can be merged to one rating D.

The probability of default of companies with ratings AAA, AA and A is equal to 0. The difference between these ratings is influenced on the financial state of companies. The companies with higher ratings have higher average profitability and liquidity ratios. Also these companies have lower average debt ratio and individual possibility of default. So according to the rating of reliable companies, a bank can determine the interest rate for loans.

Conclusions

1. This research confirmed that the discriminant analysis, logistic regression and artificial neural networks are relevant methods for the classification of banks clients. The highest classification accuracy (97%) was reached by the logistic regression model.

2. The analysis of variance, Kolmogorov-Smirnov test and factor analysis are statistical methods capable to reduce a large amount of data and help to select variables for the estimation of credit risk.

3. The percentage of credit rating criterions suggested in this research is: profitability – 50%, liquidity – 25%, leverage – 12.5%, individual possibility of default estimated by logistic regression – 12.5%. These criterions allowed to create valid rating scale for the measurement of Lithuanian companies credit risk.

4. The recalculation of classification functions coefficients when new data of clients available can improve the correct classification results of the model. Also the changing of total scores intervals can improve a model’s rating scale.

5. The described steps of the credit risk estimation model development can help other researchers to create models which allow to measure credit risk successfully.

References


Kreditų rizikos vertinimo modelio sudarymo procesas: pagrindiniai etapai ir modelio tobulinimas

Santrauka


Tyrimo tikslas – sudaryti įmonių kredito rizikos vertinimo modelį ir apsaugoti nuo sukeltų kredito rizikos. Tyrimo metu bankų klientų klasifikavimas į 7 grupes, o nepatikimi – į 1 grupę. Įmonės turi būti suskirstytos į ne mažiau kaip 7 grupes (reitingus). Todėl šiame tyrine buvo siekiama sudaryti įmonių kredito rizikos vertinimo modelį, kuriuo patikimi bankai klientai klasifikuojami į 7 grupes, o nepatikimi – į 1 grupę. Šio tyrimo objektas – įmonių kredito rizikos vertinimo modeliai.

Tyrimo metodai:

- Mokslinių publikacijų, susijusių su kreditų rizikos vertinimu, analizė.
- Įmonių aplinkos pokyčiai. Atlikus tyrimą, įmonės kredito rizikos vertinimas ir kredito rizikos vertinimo modeliai, yra teisingi. Tai ir sudaryti įmonių kredito rizikos vertinimo modelį.

Įmonių kredito rizikos vertinimo modelio sudarymas. Įmonių aplinkos pokyčiai daro įtaką įmonių kredito rizikos vertinimui. Šiuo atveju įmonių kredito rizikos vertinimo modeliai, yra teisingi. Tai ir sudaryti įmonių kredito rizikos vertinimo modelį.

Moksliniai publikacijos, susijusios su kreditų rizikos vertinimu, analizė. Įmonių kredito rizikos vertinimo modelio sudarymas. Įmonių aplinkos pokyčiai. Atlikus tyrimą, įmonės kredito rizikos vertinimas ir kredito rizikos vertinimo modeliai, yra teisingi. Tai ir sudaryti įmonių kredito rizikos vertinimo modelį.

Raktožodžiai: bankas, kredito reitingai, kredito rizika, diskriminacinė analizė.