Knowledge-Based Systems 51 (2013) 72–84

Contents lists available at SciVerse ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Feature selection in corporate credit rating prediction

Petr Hajek [a,*], Krzysztof Michalak [b]

[a] Institute of System Engineering and Informatics, Faculty of Economics and Administration, University of Pardubice, Studentská 84, Pardubice, Czech Republic
[b] Department of Information Technologies, Wrocław University of Economics, Komandorska 118/120, Wrocław, Poland

ARTICLE INFO

Article history:
Received 20 June 2012
Received in revised form 3 July 2013
Accepted 13 July 2013
Available online 19 July 2013

Keywords:
Feature selection
Credit rating
Classification
Wrapper
Mixed feature selection method

ABSTRACT

Credit rating assessment is a complicated process in which many parameters describing a company are taken into consideration and a grade is assigned, which represents the reliability of a potential client. Such assessment is expensive, because domain experts have to be employed to perform the rating. One way of lowering the costs of performing the rating is to use an automated rating procedure. In this paper, we assess several automatic classification methods for credit rating assessment. The methods presented in this paper follow a well-known paradigm of supervised machine learning, where they are first trained on a dataset representing companies with a known credibility, and then applied to companies with unknown credibility. We employed a procedure of feature selection that improved the accuracy of the ratings obtained as a result of classification. In addition, feature selection reduced the number of parameters describing a company that have to be known before the automatic rating can be performed. Wrappers performed better than filters for both US and European datasets. However, better classification performance was achieved at a cost of additional computational time. Our results also suggest that US rating methodology prefers the size of companies and market value ratios, whereas the European methodology relies more on profitability and leverage ratios.

1. Introduction

There are many instances when the credibility of a company needs to be measured. Bond investors, debt issuers, governmental officers, and companies that provide credit use credit ratings to assess the investment risk. These ratings are a basis for important decisions, and therefore there is a need for them to be as accurate as possible. Unfortunately, this usually means that many details of a company's profile have to be taken into consideration. Such a detailed analysis can be performed by experts, but this is often costly and time-consuming.

Because manual analysis of a company's profile is slow and costly, currently significant emphasis is placed on computational methods of credit rating assessment. The methods applied to this task can be broadly divided into two groups: traditional statistical methods (e.g. [36]) and artificial intelligence (AI) methods (e.g. [6,32,26]). Using traditional statistical methods is difficult because of the complexity of dependencies between various factors that influence the final rating. Nevertheless, methods such as multiple discriminant analysis (MDA) and linear regression (LR) have been applied to credit rating prediction in the literature.

AI methods are often employed when the relationships between input parameters and the outcome are too complex to describe analytically. In the case of credit rating prediction, the most promising group of methods are various classifiers that can be trained on examples – a dataset describing companies previously labeled by experts. A previously trained classifier is a model that represents dependencies between input parameters and object classification (in the case of corporate credit rating prediction – a grade representing the credibility of a company). This type of classifier has a generalization ability – it can produce ratings for yet unseen companies.

Apart from generating credit ratings for companies, corporate credit rating prediction models can be used as early warning indicators. In fact, such models may detect the start of financial crises. In addition, automatic corporate credit rating prediction can be effectively used by supervisory institutions for determining regulatory capital requirements [21].

Companies for which a credit rating is performed are described by a number of variables (features) that are used as input data for modeling. These parameters mainly reflect various aspects of economic and financial performance of a company.

An important problem in modeling credit ratings is the selection of the most appropriate set of variables. The influence of various parameters on the classification result must be quantified. Thus, only important variables can be used for the credit rating process. There is a wide variety of feature selection methods that

* Corresponding author. Tel.: +420 466 036 074; fax: +420 466 036 010.
E-mail addresses: petr.hajek@upce.cz (P. Hajek), krzysztof.michalak@ue.wroc.pl (K. Michalak).

0950-7051/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2013.07.008

can be used to select the appropriate set of variables. Sensitivity analysis (stepwise procedure) may be used to select features that have the most influence on the classification result.

In this paper a different method was used, whereby features are selected using the wrapper approach (an iterative selection procedure that selects features based on the evaluation of performance of the classifier using selected features). In addition, our research is aimed at comparing feature selection methods used for credit rating prediction. To the best of our knowledge, no previous study has attempted to compare filters and wrappers as feature selection methods within the classification process in this or related business domains. We hypothesized that wrapper approaches would contribute to a higher classification accuracy of the employed classifiers compared to filter approaches. In our research we compared filter and wrapper methods in order to verify this hypothesis.

Since two datasets were used, namely data from the US and Europe, we also address the impact of country-related determinants. Thus, our results may lead to a richer understanding of the role of individual input variables and the categories of input variables, respectively, in corporate credit rating prediction.

This paper is structured as follows. In Section 2, details of the credit rating process are presented and methods currently used in the literature for corporate credit rating prediction are described. Section 3 contains data description and a discussion of input variables used for classification. In Section 4 the feature selection process is described and four feature selection methods are introduced, namely two wrappers and two filters. The results obtained in the experiments using various classifiers are presented in Section 5. Section 6 concludes the paper.

2. Credit rating prediction – literature review

Corporate credit rating is a process in which a grade $\omega \in \Omega$ from a predefined rating scale Ω is assigned to a company. Rating agencies, such as Standard & Poor's (S&P's), Moody's, and Fitch have their own rating scales. For example, the rating scale of the S&P's is $\Omega = \{AAA, AA, A, BBB, BB, B, CCC, CC, C, D\}$ – a total of 10 grades (rating classes) that are ordered from AAA, the most promising for investors, to D, the most risky one.

Prior studies on credit rating prediction vary with respect to assessed objects, input variables used, and the set of rating classes Ω. Traditional statistical methods and AI methods have been previously employed in the literature for corporate credit rating prediction.

Studies comparing traditional statistical methods showed that the most successful methods of that type are the ordered logistic regression (OLR) and ordered probit model (OPM) [38]. The two methods have outperformed other statistical methods such as linear regression (LR) and multiple discriminant analysis (MDA) [36,37]. This may be due to the fact that OLR and OPM take the ordering of rating classes into consideration.

To use statistical methods, one has to first choose a model with a predefined structure to represent observations. Then, the parameters of the model are estimated to fit the model to the observational data. The advantage of such an approach is that the models are relatively easy to explain. However, statistical models require various assumptions to be theoretically valid.

Another approach is to use AI methods. The AI methods differ from traditional statistical methods in that they allow learning the model from data [32]. The advantage of such an approach is that AI methods usually do not require specific assumptions on the distribution of data.

Using concepts from the machine learning paradigm, the problem of credit rating prediction can be formulated as a classification problem in which rating classes used by a particular rating agency are known in advance. A typical classification procedure is performed as supervised learning. This learning type requires a sample of companies that were initially assigned proper ratings. A classifier of a chosen type is first trained using this sample. Then, the trained classifier can be used to predict the ratings of previously unseen companies.

Neural networks (NNs) are commonly used in the literature for credit rating prediction. NNs were found to be significantly more accurate than traditional statistical methods in previous studies (e.g. [6]). Hajek [26] compared the performance of a variety of NNs. Radial basis function neural networks (RBF) and probabilistic neural networks (PNNs) significantly outperformed methods such as multilayer perceptron (MLP), group method of data handling (GMDH), MDA, and LR. Because of high generalization ability, support vector machines (SVMs) also produced good results in terms of classification accuracy (e.g. [32,47]). For a small proportion of labeled companies, kernel-based approaches with semi-supervised learning [29] have provided better results than supervised learning methods.

Fuzzy logic based classifiers, such as adaptive fuzzy rule based systems (AFRBs), fuzzy decision trees (FDTs) and the Wang–Mendel algorithm have been employed by Hajek [28]. The main advantage of these classifiers is the fact that the model obtained can be interpreted in terms of membership functions and fuzzy if-then rules. However, the model can be very complex in the case of credit ratings. As a result, hundreds of fuzzy if-then rules have to be generated in order to obtain an accurate prediction model.

Other AI methods used for credit rating prediction include artificial immune systems (AISs) (e.g. [13]), case-based reasoning (CBR) (e.g. [43,65,47]), evolutionary algorithms [5], and ant colony optimization [54].

Table 1 summarizes the literature on corporate credit rating prediction. In all presented studies, data used in the tests were obtained from US companies. Rating classes were provided by two rating agencies: S&P's or Moody's. Nevertheless, it would be inappropriate to compare these studies among themeselves as they are based on different datasets (the companies included might not be the same and data were obtained from different time periods, and thus they describe companies operating in different macroeconomic conditions).

In addition to studies focused on US data, some studies have explored data from other countries and rating agencies. The assessments of Korean or Japanese rating agencies have been used in the following studies. Methods such as NNs [47], SVMs [1,2], CBR [65], Bayesian networks [72,10], and hybrid methods combining AI methods [43,4,23] have been used in these studies. Furthermore, credit ratings of sub-sovereign and municipal entities have been studied recently (e.g. [20,27]).

Considering the process of feature selection, statistical tests (one-way analysis of variance [ANOVA], Kruskal–Wallis test), factor analysis, and stepwise procedure have been used previously in various combinations [32,65,43,47,42,58]. Huang et al. [32] used one-way ANOVA to find statistically significant input variables for two datasets: Korean and US. In the case of the US dataset, 14 out of 19 features were selected with a $P < 0.1$. In particular, liquidity ratios did not have a significant impact with regard to the credit rating decision. Shin and Han [65] applied a two-stage feature selection process. At the first stage, 27 variables were selected using factor analysis, one-way ANOVA, and Kruskal–Wallis test (for qualitative variables). In the second stage, they used a stepwise procedure of MDA to reduce the dimensionality to 12 final input variables. A wide range of variables' categories was included in the resulting set of variables. In a similar manner, Kim and Han [43] performed one-way ANOVA and Kruskal–Wallis in the first stage and factor analysis with stepwise procedure of CBR in the second and third stage, respectively. Thus, the original set of 129