Applied Soft Computing 24 (2014) 494–499

Contents lists available at ScienceDirect

Applied Soft Computing

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

AdaBoost based bankruptcy forecasting of Korean construction companies

Junyoung Heo *, Jin Yong Yang

Department of Computer Engineering, Hansung University, Republic of Korea

ARTICLE INFO

Article history:
Received 26 November 2013
Received in revised form 29 June 2014
Accepted 1 August 2014
Available online 12 August 2014

Keywords:
Bankruptcy
AdaBoost
Z-score

ABSTRACT

A lot of bankruptcy forecasting model has been studied. Most of them uses corporate finance data and is intended for general companies. It may not appropriate for forecasting bankruptcy of construction companies which has big liquidity. It has a different capital structure, and the model to judge the financial risk of general companies can be difficult to apply the construction companies. The existing studies such as traditional *Z*-score and bankruptcy prediction using machine learning focus on the companies of non-specific industries. The characteristics of companies are not considered at all. In this paper, we showed that AdaBoost (adaptive boosting) is an appropriate model to judge the financial risk of Korean construction companies. We classified construction companies into three groups – large, middle, and small based on the capital of a company. We analyzed the predictive ability of the AdaBoost and other algorithms for each group of companies. The experimental results showed that the AdaBoost has more predictive power than others, especially for the large group of companies that has the capital more than 50 billion won.

1. Introduction

In recent times, bankruptcy of construction companies has rapidly increased due to the recession in the real estate sector. Since 2008, a continuously decreasing profit ratio in the construction industry has had a negative influence on the liquidity of construction companies. According to the economic outlook report of 2013, it is predicted that the liquidity crisis experienced by construction companies will persist due to the continuing recession in the housing construction business [1]. While the construction industry has bigger social ripple effects caused by bankruptcy than other industries, due to the nature of its capital structure, debt-to-equity ratio and cash flow are different from other industries, bankruptcy predictions for construction corporations become more difficult.

The ratio of the construction industry to GDP in Korea is 5.9%, which is higher than the average ratio of other OECD countries, 5.1%. Investment in construction attempts to increase economic growth with large amounts of capital. While the USA shows a notable recovery trend, Korea is hindered by a significant decrease in investment in the construction industry due to the recession in the real estate business.

The construction industry has high leverage and debt-to-equity ratios. Positive cash flow in the construction business is highly concentrated in the latter parts of the projects. Construction companies are highly sensitive to economic cycles and bankruptcy rapidly increases in an economic downturn. Because the construction industry requires high leverage, increased bankruptcies of construction companies will be a big burden to creditor banks which provided the construction loans. Nevertheless, bankruptcy forecasting models mainly focused on financial institutions, and construction-specific studies have rarely been carried out.

Increases in bankruptcies of construction companies create a big burden on the banks that provide loans to the companies. In fact, some savings banks that were primarily dedicated to real estate project financing went bankrupt recently along with the bankruptcy of their borrowers, the construction companies. Not only savings banks, but also commercial banks could not avoid the impact. These banks should secure significant amounts of capital to cover losses for the possible risk of borrower bankruptcy. As the construction business is entwined in a complicated system involving numerous subcontractors, bankruptcy of a single construction company triggers a chain reaction of bankruptcies of other companies. Furthermore, since the construction industry is closely related to raw material industries such as cement and steel, its recession has huge influence on other industries. The construction industry also has a big employment-inducing effect, its bankruptcy has greatly affected ordinary households. Notwithstanding the fact, bankruptcy forecasting models so far have mainly focused on

* Corresponding author at: 116 Samseongyoro-16gil, Seongbuk-gu, Seoul 136-792, Republic of Korea. Tel.: +82 2 760 8039; fax: +82 2 760 4346.
E-mail addresses: jyheo@hansung.ac.kr, jyheo0@gmail.com (J. Heo), jyang0112@gmail.com (J.Y. Yang).

http://dx.doi.org/10.1016/j.asoc.2014.08.009
1568-4946/

financial institutions, and few studies are conducted with construction industry-specified models.

Studies of bankruptcy forecasting models based on financial statements of a company have been conducted in diverse ways for long time. The subjects of the model, however, were the general firms, and the models may not be proper for accurately forecasting companies having large liquidity issues such as construction companies. The construction industry is a capital-intensive industry that requires long-term project periods, huge investment, and takes a long time to receive returns from the investment. Therefore, it has a different capital structure from other industries, and the same criteria used for other industries cannot be applied to effectively evaluate its financial risk [2].

Altman's *Z*-score has been commonly used as a bankruptcy forecasting model [3] so far. *Z*-score was first published in 1968 and it forecasted the likelihood of a company going bankrupt by using a simple formula. *Z*-score classified the results into three categories and evaluated the corporate status as dangerous, moderate or safe. When a company falls in the "dangerous" category, it has a high likelihood of bankruptcy within two years, while in the "safe" category, it has a low likelihood of bankruptcy. When the company falls in the "moderate" category, it is not easy to forecast the risk. Many of the construction cases in this study fell in the "moderate" category, which made it hard for us to forecast its risk.

Along with the development of machine learning by using a computer, studies of forecasting corporate bankruptcy have been active recently based on machine learning. Pattern recognition, which is a representative application area of machine learning, is applied in forecasting corporate bankruptcy. Patterns were analyzed based on financial information of a company and then we judged whether the pattern belongs to the bankruptcy risk group or safe group. The representative machine learning models used in the bankruptcy forecasting are Artificial Neural Networks [13,14,7] and Adaptive Boosting (AdaBoost) [2]. Some research used Support Vector Machine (SVM) [15,16,5,6]. There are also diverse hybrid studies that combined these models [17,8]. All the previous work used financial statements as an input in their forecasting models. These works are characterized by the machine learning algorithm and company's country. They all classified the financial statements as capital, assets, sales, income, and liability. The data was transformed into ratio because the scale is very different for each company. For example, the liability data was transformed into total liability to total assets.

Previous studies of bankruptcy forecasting models using *Z*-score and machine learning were carried out using general companies. Therefore, the specific characteristics of each industry have not been considered. This study compared the predictive ability of each model according to the size of the construction companies and verified that Adaptive Boosting has the best predictive ability.

This study classified construction companies into three classes according to the size of capital. It analyzed predictive ability of the AdaBoost for each class. For relative performance measurement of AdaBoost, we conducted comparative analysis with other machine learning models such as Artificial Neural Networks, SVM and Decision Tree. Predictive ability of each model was measured with financial data from the companies that went bankrupt in the period from 2008 to 2012 with those in normal operation as of 2012. As a result, it was found that AdaBoost had good predictive ability compared to other models. In particular, for companies of which capital was more than 50 billion won, it showed outstanding predictive ability.

The contribution of our work is that we have selected 12 variables for the model of bankruptcy forecasting and come to the conclusion that AdaBoost with a decision tree is more appropriate for the model that other machine learning algorithms, especially for Korean construction companies.

The structure of this study is as follows. Section "Description of the research issues" explains the issue of bankruptcy forecasting and Section "Descriptions on the data" describes the data of construction companies used in this study. Section "Adaptive Boosting" explains AdaBoost, the major machine learning method of this study. Section "Experiment" shows the experiment results and final section contains the conclusion.

2. Description of the research issues

Forecasting bankruptcy of a company is an important issue in business management. The goal is to discern sound companies from the companies having a likelihood of going bankrupt. That is, its goal is to construct a risk forecasting model of a company and make proper decisions accordingly based on the predictions. In order to forecast bankruptcy risk of a company, it is essential to know the current financial status of the company. As this information may have great difference in values depending on the size of the company, ratios are usually used in bankruptcy forecasting models [4]. For example, to indicate current assets, the ratio obtained from dividing by the total assets is used.

Although everybody acknowledges the importance of bankruptcy forecasting models, there are different opinions regarding the point in time a company is judged to be in a bankrupt state. In this study, all the cases of workout, receivership, and bankruptcy are considered as bankruptcy. When a company is in financial hardship, many stakeholders including shareholders, creditors, and employees face significant difficulties. Since all three types of financial hardship of a company can create huge damage and loss to its economic stakeholders, it is considered reasonable to define companies in such state as bankrupt.

Diverse kinds of financial data variables can be used for a bankruptcy forecasting model. Forecasting is an act which attempts to predict the state of the future based on current data. In this study, we intend to forecast the likelihood of bankruptcy of a company based on one year of its financial data. In order to set a model and verify the predictive ability of the model, this study used the financial data from bankrupt companies one year prior to the point of bankruptcy and that of a normal company for comparison.

3. Descriptions on the data

The financial data to be applied to the bankruptcy forecasting models is based on the financial statements of the companies for the past 5 years. We used the data from 5-year time periods regardless of the actual time because we need to exclude the effect of the economy in a given year. This study extracted the following financial data from the NICE DnB [9] which retains the financial data for all construction companies in Korea.

We classified companies into two categories: bankrupt and normal. The bankrupt companies are those that went into workout, receivership, or bankrupt during the period from 2008 to 2012. The normal companies are those that were not in bankrupt state as of December 2012. Eventually, we selected 1381 bankrupt companies and 28,481 normal companies. Then we collected the financial data of the bankrupt companies and normal companies. As for the financial data of the bankrupt companies, we used the financial data one year prior to bankruptcy. As for the financial data of normal companies, we used their 2011 financial data. That is, we applied the financial data one year prior to bankruptcy for bankrupt companies and the financial data of the previous year of the normal companies to the forecast models to see the level of bankruptcy after one year had passed.

Based on the collected financial data, we produced the following variables and applied them to the model.