



A practical approach to bankruptcy prediction for small businesses: Substituting the unavailable financial data for credit card sales information

Jong Sik Yoon^a, Young S. Kwon^{b,*}

^a Solution Business Team, D&B Korea, Credit Center Bldg., Ahyeon-dong, Mapo-gu, Seoul 121-709, South Korea

^b Dongguk Univ-Seoul, Department of Industrial & Systems Engineering, Seoul 100-715, South Korea

ARTICLE INFO

Keywords:

Bankruptcy prediction
Small business
Support vector machines
Data mining
Credit card sales

ABSTRACT

Small businesses are open to the elements of both consumer and business credit risks. One of the problems in bankruptcy prediction for small businesses is that the official financial data in most cases are not available for evaluating the business credit risks. In order to alleviate this problem, we propose to use the credit card sales information as a substitute for the official financial data in developing a bankruptcy prediction model. In most cases, the credit card sales information is available because most small businesses are member stores of credit card processors. First, we derived several variables using the credit card sales information, including business period, sales scale, sales fluctuation, sales pattern and business category's bankruptcy ratio, etc. Then we developed support vector machines (SVM) based model. The empirical analyses show that credit card sales information is an acceptable substitute for the financial data in predicting the bankruptcy of small businesses. In addition, the proposed SVM model exhibits superior performance compared to other classifiers such as neural networks, CART, C5.0, multivariate discriminant analysis (MDA), and logistic regression analysis (LRA).

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Since the financial crisis in 1997, the Korean government strengthened the financial support for the start up of small businesses in order to promote the economy. As the financial institutions were encouraged to make loans to the start ups, the need for the appropriate models for the credit evaluation and bankruptcy prediction has been keenly felt. However, as small businesses are open to the elements of both consumer and business credit risks, it is hard to develop a model which reflects both features. Furthermore, the model is even harder to develop due to the difficulty in obtaining objective criteria for the business properties such as financial information, income statements, predictive revenue, location and business potential, etc.

As the Korean government has introduced a variety of policies to expand credit card usage in late 1990s to promote the economy, the issuing number of credit card and the number of merchants in Korea has rapidly increased as in Fig. 1. Therefore, many of small businesses gained much revenue through credit cards usage and sales volume could be estimated through the credit cards slips (Yoon, Kwon, & Roh, 2007).

The purpose of this paper is to develop a bankruptcy prediction model for small businesses by utilizing credit card sales informa-

tion instead of using the insufficient financial data for small businesses. We thereby derived a variety of variables that is highly related to credit card sales information, and adopted a machine learning technique of support vector machines (SVM). The rest of this paper is organized as follows. In Section 2, we review previous researches on the bankruptcy prediction models for small businesses. Section 3 describes the data, and Section 4 presents the model and experimental results, followed by the implication and the conclusion.

2. Previous researches

The bankruptcy prediction has been a long standing research issue. There have been a lot of researches and models for bankruptcy prediction. Statistical methods such as multivariate discriminant analysis, regression and logistic regression have been widely used in bankruptcy research (Gentry, Newbold, & Whitford, 1985; Ohlson, 1980). Since the 1990s, many researches have showed that machine learning approaches of artificial neural network, decision tree and SVM performed well in some cases (Huang, Chen, Hsu, Chen, & Wu, 2004; Kwon, Han, & Lee, 1997; Tam & Kiang, 1992).

However, research on bankruptcy prediction for small businesses has not attracted the international attentions so far, because of the unavailability of financial data and the international differences in the definition and policies for small businesses. Since late 1990s, there has been an upward trend in running one's own

* Corresponding author. Tel.: +82 2 2260 3375; fax: +82 22269 2212.
E-mail address: yskwon@dongguk.edu (Y.S. Kwon).

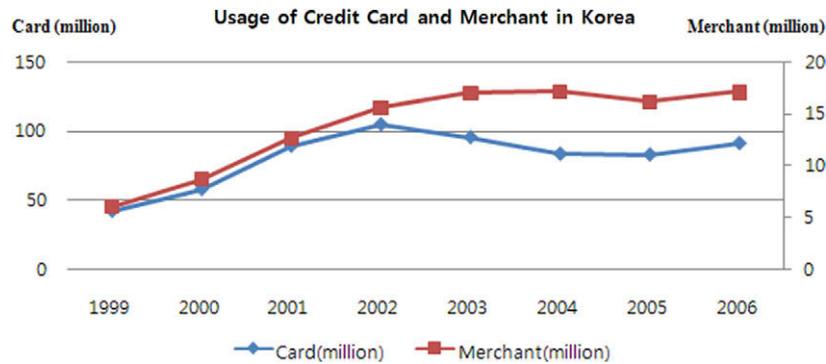


Fig. 1. The number of credit card and merchants in Korea.

business in many countries including Korea. In 2008, the Korea National Statistical Office announced that the number of small business were about three million which was almost 90% of all enterprise in Korea and the number of employees in small businesses was about six million, which consisted of almost 40% of all employees in Korea. Such figures imply the importance of small businesses.

The followings are some previous researches which relates to small businesses. Altman, Haldeman, and Naraynan (1977) indicated that a critical component in a business evaluation is its size, and the rate of bankruptcy is inversely correlated with company size. They stressed on the need for further research on the credit evaluation of small businesses.

Berger, Frame, and Miller (2005) also stated that the application of credit scoring to small businesses may reduce the economic risk. They noted the importance of forming a confident relationship between small businesses and financial agencies. Frame, Scott, and Woosley (2004) suggested that lending agencies had shunned lending to small businesses because it is not easy to evaluate the small businesses' risk because of lacks of financial information. They stated that using the suitable model for small enterprises was to release the unsymmetrical information between creditors and debtors, in order to increase the lending to small enterprises. They particularly asserted the usefulness of credit evaluation of small enterprises. They showed that such a system was more effective for low-profit companies, which faced difficulties in securing funds, rather than profitable companies.

Gim and Jong (1999) studied a model capable of predicting the bankruptcy of small business through MDA, while Gim, Kim, and Lee (2005) and Kim (2006) studied models to predict insolvent by utilizing logit analysis, combined with AHP for incorporating non-financial information of small business. Due to the lack of financial data based on the model for credit evaluation of the company, they initially used the AHP technique by utilizing the bank managers' know-how to make lending assessment decision for small businesses, using the non-financial information.

3. Research data

In this study, the credit card sales information of small business merchants is used in lieu of insufficient financial data. In Korea, as each credit card company only manages its own cards, they do not have information on other credit card sales. However, a credit card processor which has a responsibility of managing information of every transaction gathered at the terminal unit of all member stores, and then sends the information to the affiliated credit card provider. All transaction data are controlled by credit card processors. Credit card processors also provide credit card sales information to the affiliated credit card provider at any time in agreement with the merchants.

We collected the monthly credit card sales data for 28 months between January of 2000 and April of 2002, for 412,773 merchants affiliated with the K-VAN service company. The delinquency information of merchants obtained from the Korea Federation of Banks was matched with the credit card sales data. Among 412,773 merchants, we deleted the data on unmatched on merchants between the K-VAN service data and credit card provider data and the missing values. Other variables were derived by analyzing business type, continual transaction terms, average transaction value, standard deviation of transaction, and transaction pattern, etc. These variables were used as input variables in the model. Among 412,773 cases, we have about 3% of bankruptcy cases. To avoid the overfitting, 5000 samples were randomly selected each from the bankruptcy and the non-bankruptcy stores. Among them, 7000 samples were selected for learning and the remaining 3000 samples were used for the test.

4. The proposed model and empirical analyses

4.1. Definition of bankruptcy in small business and the derivation of input variables

It is not easy to collect the information on small business bankruptcy because small business does not usually report its business closure or bankruptcy to the relevant financial institutions. Therefore, we define the bankruptcy of small businesses in terms of delinquency and the suspension of card transaction because the suspension of card transaction implies the closure of business.

In order to define the bankruptcy of small business, we utilize the information of the suspension of card transactions at merchants and the delinquency of merchants which are obtained from the Korea Federation of Banks. We found that the number of cases of suspended credit card transaction before the month in which the delinquency was identified were more than those suspended in the month after the identification, as shown in Fig. 2. Especially, the suspended cases within 6 months before delinquency and 3 months after delinquency were heavily populated. Therefore, we defined these cases as the bankruptcy.

To build the model, we derived various input variables using the merchant's credit card sales information.

First, we derived 24 variables which are related to credit card transaction period, number of transactions, and transaction amount by using credit card sales information. Among them, for the variables of the month suspended, average of sales, maximum and minimum amount of sales, variance of sales, average of transaction and variance of transaction during 3 months and 6 months period, *t*-test statistics for the 6 months period are higher than those of 3 months period, so the 13 variables including the derived variables of 6 months period are selected for the input variables as in Table 1.

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات