Applying a nonparametric random forest algorithm to assess the credit risk of the energy industry in China

Lingxiao Tang, Fei Cai, Yao Ouyang

Hunan Normal University, China  
Changsha University of Science & Technology, China

ARTICLE INFO

Keywords: Random forest  
Credit risk  
Energy industry  
Overdraft ratio

ABSTRACT

With the rapid growth of the credit card business in China’s energy industry, credit risk is gradually revealed. This study aims to scientifically measure the credit risk of credit cards used in China’s energy industry and to lay the foundation for comprehensive credit risk management. Based on an analysis of the factors influencing credit risk influencing factors, this study applies the random forest algorithm and the monthly data of credit cards used by energy industry customers in a branch of the Postal Savings Bank of China from April 2014 to June 2017 to build an effective credit risk assessment model and scientifically measure the credit risk in China’s energy industry. The results suggest that credit card features like the overdraft ratio and the amount of credit card expenses within a month have significant impacts on credit risk, our model’s comprehensive prediction accuracy is as high as 91.5%, and its stability is satisfying. These findings can provide valuable information to help banks improve their credit risk management.

1. Introduction

A credit card system is an innovative financial product with various advantages; for example, such systems are easy to carry and use, permit credit loans and cash withdrawals, and allow payments in installments. Since 1915, when the credit card system was introduced in the United States, credit cards have gained enormous popularity and immediately became the most dynamic credit tool used to pay for goods. In China, the development of the credit card industry was relatively late. The first credit card was issued in June 1985 by the Bank of China’s Zhujiang branch, although the development of the industry over the subsequent decade remained slow. It was not until 2003 (referred to as the "first year of the credit card") that the credit card industry began to rapidly develop and, thereafter, maintain high and rapid growth. According to statistics published by the People’s Bank of China, by the end of 2016, the number of in-use credit and debit-credit cards reached 465 million, with a per capita credit card ratio of 0.31 and reported increases of 7.64% and 6.27% compared to the previous year's figures. The rapid development of the credit card industry has generated significant profits for China’s banks. According to the annual report of China CITIC Bank, in 2016, the net profit generated by the credit card industry increased by 77%. In addition, the net non-interest income of the industry was RMB 16.886 billion, thus accounting for 72.82% of the non-interest net income of retail bank branches and 35.4% of the banks’ total non-interest net income. However, high growth and profits tend to lead to numerous potential financial risks, such as credit, liquidity, and operational risks. In fact, the loss incurred by credit risk is one of the most important risks in the credit card industry and, consequently, has gained much attention. In 2013, the total outstanding credit overdue by six months was RMB 25.192 billion. In the same year, the growth rate of outstanding credit was 71.86%, which was higher than that of credit cards in use (18.03%). In 2015, the number of in-use credit cards encountered negative growth for the first time; however, outstanding credit overdue by six months still reported growth of 6.33%. In 2016, the growth rate of the total outstanding credit overdue by six months reached 40.86% (RMB 53.568 billion), which was four times greater than the growth rate of credit card usage.

This study is sponsored by the “Study on The BRICS Contingent Reserve Arrangement from the Perspective of International Lender of Last Resort” (16AJL011) (key project of the National Social Science Foundation of China); “Model and Empirical Study on Economic Growth of Large State Based on Scale Advantages” (71373075) and “Model and Empirical Study for Large Countries to Realize their Factor’s Supply and Demand Equilibrium” (71373075) (projects of National Natural Science Foundation of China); and the Hunan Province Graduation Student Research and Innovation Project (CX2017B483). The authors hereby express thanks to the anonymous reviewers for their constructive suggestions.

E-mail address: ouyangyao@yeah.net (Y. Ouyang).

Data are compiled from the Overall Situation of Payment Systems in 2016, issued by the People's Bank of China.

Data source: annual reports issued by China CITIC bank.

https://doi.org/10.1016/j.techfore.2018.03.007

Received 15 October 2017; Received in revised form 6 January 2018; Accepted 6 March 2018

0040-1625/ © 2018 Elsevier Inc. All rights reserved.

Please cite this article as: Tang, L., Technological Forecasting & Social Change (2018), https://doi.org/10.1016/j.techfore.2018.03.007
than the growth in the number of in-use credit cards.\(^3\)

Superficially, this major credit risk in China appears to be caused by the lowering of customer assessment criteria and the excessive issuance of credit cards owing to market share competitiveness among banks during the early development stage. However, from an information economics perspective, the actual causes are moral hazard and adverse selection as a result of information asymmetry. Banks are generally not fully aware of applicants' credit statuses before issuing credit cards and, thus, tend to apply self-defined criteria established on the basis of the average credit status in society to determine customers' illegibility. In addition, standard measures to reduce credit risk include acquiring applicants' income or property statements, implementing high interest rates, and limiting the maximum amount of credit granted (credit limit). These measures, however, are not well accepted by certain high-quality, low-risk customers, who consequently withdraw from the market. As a result, the market has become dominated by low-quality, high-risk customers in urgent need of money. This phenomenon is known as reverse selection. Furthermore, it is difficult for issuing banks to remain updated on the customers' economic conditions once the credit cards have been issued or on their economic activities following each cash withdrawal; for instance, some cardholders may use the credited amount to engage in high-risk economic activities, thereby leading to a moral hazard. Evidently, the lack of social integrity, absence of a well-established credit information system and comprehensive legal system, and under-developed information technology systems in China have led to common and serious problems of moral hazards and adverse selection. In response to the growing financial risks and fragility of the financial system in China, in 2016, the Central Economic Work Conference stressed that "the prevention and control of financial risk should be placed in a more important position." Thus, studying the credit risk of China's credit card systems has gained considerable significance. A fundamental pillar in China's economy is the energy industry, and its healthy development is vital to the nation's economic development. Businesses in the energy industry, particularly the new energy industry, rely on bank loans as a major capital source. However, the large-scale bankruptcy in the new energy industry has affected the stability of China's financial system (Xu and Fan, 2014). Thus, adopting credit management techniques and advanced technology (e.g., data mining) from other countries to identify, monitor, and manage credit risk within the credit card system has become a prominent topic in the context of China's banking industry.

This study aims to scientifically measure the credit risk of credit cards in China's energy industry and to lay a foundation for comprehensive credit risk management. Based on an analysis of the factors influencing credit risk, this study applies a random forest (RF) algorithm to build an effective credit risk assessment model and tests it on monthly data from credit cards used by energy industry customers in a branch of the Postal Savings Bank of China from April 2014 to June 2017, and, thus, it scientifically measures the credit risk of credit cards used in China's energy industry. The results show that features like the overdraft ratio and the amount of credit card expenses within a month have significant impacts on credit risks. The model's comprehensive prediction accuracy is as high as 91.5%, with satisfying stability.

This study has the following academic values and contributions. First, the research object is novel. Existing studies of credit risk measurement mainly focus on the real estate, construction, or steel industries, with little attention paid to the energy industry. However, the energy industry is a fundamental industry in China's economy. Credit risks in this area have gradually emerged in recent years and have become an unignorable systemic risk factor. Therefore, research in this area has significant practical meaning. Second, our research method performs better than previous research methods do. Previous studies on credit risk mainly built their credit risk assessment models using multivariate discriminant analysis, logistic regression, neural networks, and support vector machines. However, these methods have many shortcomings. For example, multivariate discriminant analysis requires the sample data to be subject have a multivariate normal distribution and equivalent covariances; logistic regression is prone to multiple collinearity; neural networks have ambiguous structures; and support vector machines require symmetric sample data, which is hard to obtain in China due to the lack of a comprehensive and effective credit risk database. However, the RF algorithm adopted in this study not only has no distribution data scale requirements, but it also can overcome multi-collinearity. In addition, it is more tolerant to outliers and noise, and is less likely to have overfitting issues. Thus in view of the above analysis and the possible correlations between the features in this study, the RF method is more appropriate in this paper. Third, we include overdraft ratio characteristics in the credit risk factor system for the first time. When analyzing the factors influencing credit risk, we use the mean decrease accuracy and mean decrease Gini methods to rank the importance of all the features. The results show that the overdraft ratio has the most significant impact on credit risk, so we include the previously ignored overdraft ratio characteristics in the influencing factor system for the first time. Fourth, our constructed credit evaluation model has a high prediction accuracy. Due to the inherent advantages of the RF algorithm and the extracted features, such as the overdraft ratio, our credit risk assessment model has a high accuracy and satisfying stability in comprehensive prediction and can be used as a valuable resource for banks aiming to improve their credit risk management.

The structure of this paper is as follows. Section 1 describes the status of credit card credit risk in China. Section 2 reviews the development of credit risk assessment models for commercial banks, analyzing the shortcomings of each model and the advantages of the RF algorithm. In Section 3, we construct the proposed credit risk assessment model on the basis of the RF algorithm and, then, use empirical data to examine the applicability of the proposed model. More specifically, we elaborate on the general principles of the RF algorithm, introduce research data, and analyze factors influencing the credit risk of the credit card system. In addition to the variables used in past research, this study includes several new variables in the feature set, such as an overdraft ratio, the value of the first credit card expense, and credit limits. Further, we employ the mean decrease accuracy and mean decrease Gini methods to screen the proposed feature set. Finally, we construct the assessment model on the basis of these key features and introduce evaluation indicators targeting unbalanced data (default and non-default samples) to assess the potential models and identify an optimal model. Section 4 concludes with key findings and future research directions.

2. Review of credit-risk assessment models for commercial banks

This study adopts the narrow definition of credit risk, which refers to credit card issuers' possibility of loss owing to unsecured revolving credit and characteristics such as unscheduled loans, the lack of a fixed location, the large number of borrowers, and the relatively small amounts of individual loans. Credit card risk mainly includes credit, fraud, and operational risk. A study conducted by the World Bank on the global banking crisis of 1992 suggests that credit risk is a common cause of bankruptcy among banks. Further, China's indirect financing structure renders credit risk the most serious issue confronting the financial industry. Therefore, constructing an assessment model to effectively identify credit risk is of significance to both banks and supervising authorities. Next, this article describes several commonly used credit risk assessment methods and provides a brief review.

---

\(^3\) Data source: Overall Situation of Payment Systems in 2015 and 2016, issued by the People's Bank of China.
دریافت فوری متن کامل مقاله

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