The Applications of Big Data Analysis in the Credit Business of Commercial Banks

Huilin Xie *

Department of Rotman Commerce University of Toronto, Toronto, Canada

*Corresponding author: huilin.xie@mail.utoronto.ca

Abstract. As big data technology becomes more and more mature, its applications in the finance industry become more widespread. Big data technology can address some issues in the traditional credit business in commercial banks and expand the scope of business. This study focuses on the characteristics of big data analysis in the credit business of commercial banks and the corresponding strategies of risk management. To be specific, this paper summarizes current studies on the topic through careful analysis and points out advice for improvement in big data analysis in the credit business. According to the analysis, big data techniques can play a role in target marketing and the development of customized services based on customers' images. Furthermore, big data applications can significantly reduce manual labor. In addition, big data technology helps in early identification of risks and risk control since banks can know the borrowers better through it. These results shed light on guiding further exploration of implementation big data analysis to bring competitive advantages to commercial banks.

Keywords: Bid data; Commercial banks; Credit business.

1. Introduction

Big data has been a widely-discussed topic in society in recent years. Big data is characterized by a large amount of data, diverse sources, rapid growth, fast generation and processing, low value of the data, authenticity and accuracy, as well as complexity [1]. Its processing requires new technology to deal with massive, classified information in a given period, which is more efficient rather than the conventional method [2]. Rather than analyzing randomly selected data, this technology analyzes all data of various types quickly. Traditional data processing usually utilizes more specific types of information and analyzes it by random sampling. Traditional data analyzing methods often overlook some trivial information and cannot precisely confirm the standard format. These kinds of information are crucial for data analysis, especially if there is a large amount of data. Big data technologies can address the problems to some extent. Besides, big data can carry out forward analysis based on visualization analysis and data mining with more predictive value. Meanwhile, conventional data analysis is always "backward analysis" that summarizes things already happened.

Big data analysis has provided directions for improvement in many industries. The banking industry mainly employs traditional database and tool analysis. The applications of big data have begun to take shape since 2010. At that time, the world just entered the ZB (Zettabyte) era and maintained a 58% annual growth rate in data volume [1]. In 2014, big data technologies had been applied to the financial sector [3]. The development of the distributed file system, distributed database and other technologies enhanced the information processing capability of commercial banks [4]. As information technology and financial technology grow rapidly, the applications of big data had been relatively fully developed and utilized in a wide range of industries in 2016 [5]. The fast growth of data has become the driving force for promoting the continuous innovation of information, and the operation mode in many industries has changed quietly. In the era of big data, various industries will face more challenges and opportunities. This will bring a new stage to the modern information society.

In addition, In the era of Internet Finance, the flaws of traditional banking business got more exposure, e.g., information asymmetry, high costs of borrowers' information auditing, low efficiency. These problems impede the operation of the credit business, one of the main businesses, of commercial banks. It is difficult for small and medium-sized enterprises and individuals to collect

enough collateral and hard information required by the traditional credit business model, therefore difficult for them to obtain loans.

This paper mainly adopts the literature research method, analyzing and summarizing the existing studies of applications of big data techniques in the credit business of commercial banks. Combined with theories and cases, the paper discusses the main applications of big data analysis in the credit business and its benefits over the traditional credit business model of commercial banks. It also provides corresponding suggestions for improvement. The remainder of the paper is arranged as follows. The Sec. 2 will introduce the information asymmetry theory and the long tail theory. The Sec. 3 will discuss the current issues in the credit business of commercial banks and three main applications of big data analysis to solve these problems. The Sec. 4 will raise two suggestions for employing big data technology for commercial banks: improving data acquisition and management as well as well-performing professional personnel training. Finally, the Sec. 5 summarizes the whole paper and describes its significance for other researchers.

2. Theories Related to the Credit Business of Commercial Banks

2.1 Information Asymmetry Theory

In the 1970s, George Akerlof first raised the phenomenon of information asymmetry through the observation of the market for second-hand cars. Second-hand car sellers and buyers have different information and sellers know the specific conditions of the car more clearly, which can lead to shoddy goods flush and bad money driving out good money. Owners of poor-quality cars can offer lower prices than high-quality car owners, and unwitting consumers would buy second-hand cars of lower prices. This can eventually result in poor-quality cars taking a greater share of the market and crowding out the high-quality goods, as well as the average quality of used cars in the market drops.

In the area of credit business, this phenomenon also has non-negligible impacts. Customers of commercial banks, also known as loan borrowers (enterprises and individuals) in seek of funds, have a deeper understanding of their financial situation and ability to repay the loans. However, loan lenders, the commercial banks themselves, often face obstacles in knowing the actual positions of the borrowers. Despite banks asking for financial information and records of borrowers' loans and credits, borrowers may falsify this information, resulting in adverse selection in the credit market. Even if borrowers provide true information about themselves, they may still use the loan for unexpected purposes after obtaining it, including purposes in line with their individual interests but can cause a moral hazard in the market. As a result of adverse selection and moral hazard, commercial banks are facing an increasing number of bad loans [6].

2.2 Long Tail Theory

According to the long tail theory, although the sales volume and market share of non-best-selling products are small, the various kinds and low operating costs make their total profit roughly equivalent to or even more than the profit of best-selling products. The theory is contrary to the traditional “80-20 rule”. By plotting the overall profit of the company (seen from Fig. 1), the meaning of the long tail theory can be seen as more straightforward. The X-axis displays the type of goods, and the Y-axis shows the volume of goods. The head represents of profit of best-selling products, which is short in volume but high in overall profit. The tail represents the profit of the non-best-selling products. Though the individual profit is lower, the long tail allows non-best-sellers to make a significant profit on the whole. The long tail theory emphasizes the importance of small customers, reflecting the huge potential market demand for non-best-selling products.

The implication of this theory for the credit business of commercial banks is that banks should pay attention not only to medium-sized and large companies but also to small and micro businesses as well as individuals. Commercial banks need to improve their ability to collect and process soft information to attract a large number of small customers, thus expanding their business scope and increasing their profit [7].

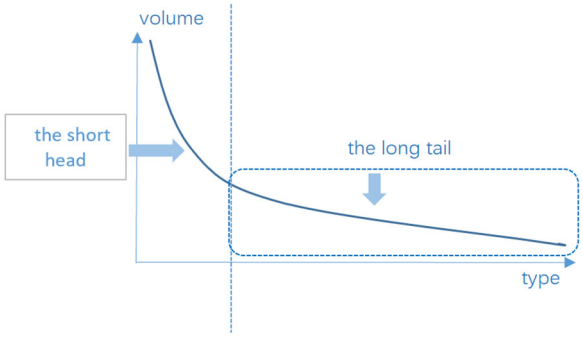

Fig. 1 The long tail theory.

3. Applications of Big Data in the Credit Business of Commercial Banks

3.1 Current Issues

Primarily, the acceptance of the credit services of commercial banks is low among small and micro businesses. According to the Report of Development of China's Small and Micro Businesses 2014, only less than 60% of small and macro businesses in need of funds would finance through bank loans. Although the pass rate of those businesses applying for credit loans is not low, many firms are reluctant to apply for loans from banks due to subjective reasons. They think of the review process as complex and the information preparation as difficult and they are unfamiliar with the review steps.

In addition, official information (e.g., enterprise financial statements) is always necessary for the information audition in the traditional credit business of commercial banks. Small and micro enterprises, however, often face issues (e.g., their financial statements not following the standard format, inadequate official information disclosure, and lack of collateral), which is different from the formal management and information collection of medium-sized and large companies. The inadequate and badly organized information brings trouble for the audition of commercial banks. Those kinds of information have high costs of review. Given each individual small and micro business usually only requires a limited number of loans, commercial banks must expand the business scope of credit loans to promote growth in the overall operation and profit scale, leading to further increases in the total costs. Meanwhile, banks cannot achieve economies of scale due to the big differences in individuals and the inability to summarize a standard model. The high operating costs for small borrowers would cause commercial banks more inclined to give loans to large companies with good credit. Therefore, the obstacles for small and micro firms to gain credit loans would be overlooked and issues would persist.

Another issue is with risk control, including risk identification, elimination, and recovery of losses. Traditional credit businesses can identify and control risks associated with companies of a certain scale according to previous procedures and regulations. The scattered distribution and large volume of small and micro businesses and individuals make it difficult and costly for commercial banks to identify and trace risks. Small and micro businesses have low default costs and can easily avoid repaying loans by transferring assets in cases of poor operation. Hence, it is more difficult for commercial banks to control risks and recoup losses, which is one of the reasons that discourages them from lending to small customers and makes borrowing money more difficult for them.

3.2 Applications of Big Data Analysis in the Credit Business of Commercial Banks

The traditional marketing model is too limited. Big data technology can help collect information about the target customer group. Banks can accurately describe their target customers in terms of characteristics such as the purpose and amount of planned loans and the firms' main operation. It will enable them to choose more suitable marketing channels and marketing strategies. Commercial banks can work with institutions that have data resources to carry out targeted marketing along with mass

marketing. By applying big data analysis to depict the images of customers, commercial banks can better understand their needs and develop appropriate new services, thereby cutting development costs. The accumulation of information and experience during the process is likely to enable commercial banks to build a virtuous circle in the field of big data optimization of credit business.

In the traditional bank credit business, a vast amount of staff is involved in the processes of document submission and audition as well as subsequent risk identification and control. This would not only lead to low efficiency and high cost but also various mistakes and increased difficulty in information management. The application of big data analysis greatly eliminates manual labor. Big data technologies can open and expand channels for information collection, utilize soft information from multi-platform systems, and offer effective support for data submitted by small customers. As a result, big data techniques can be used effectively to identify customers and reduce the cost of investigation and verification. In addition, the ability of big data to analyze information also reduces time and management costs quickly and automatically.

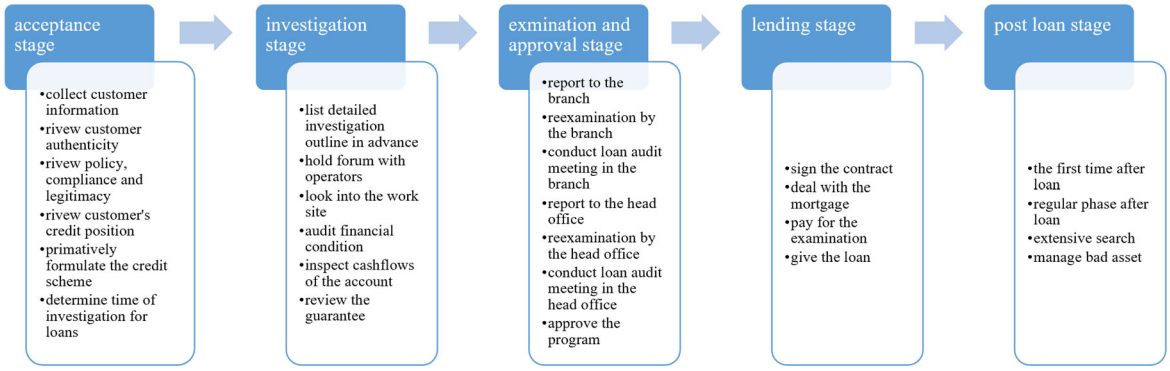

Fig. 2 Flowchart of loan management of commercial banks

Conventional commercial bank loan management contains five stages, a total of over 20 steps, as shown in Fig. 2 [8]. These processes are too complicated for small and micro enterprises and individuals, thus undermining some potential customers' willingness to get loans. By contrast, big data technology uses a large amount of information collected from other channels as a supplement to assist analysis, allowing flexible adjustments to the information submission requirements and simplifying the procedures by reducing manual intervention. Credit levels and repayment ability of borrowers can therefore be assured to meet standards, and the loan process can be handled more conveniently and efficiently. There are practical applications for big data in risk warning and control, as well. Compared to human operation, big data technology can respond to risks brought by changes in a timelier manner because of its massive volume of data collection and the capability to process it quickly. Big data monitoring of the process and loan enterprises can recognize the points of risk occurrence and give early warning as soon as possible, thereby giving commercial banks the opportunity to take measures for adjustment and prevent the risk from further expanding.

It is critical for commercial banks to establish early risk warning and monitoring departments using big data technology. First, they should collect comprehensive customer information, then rate it to ensure its authenticity and validity, followed by filtering and analyzing the information. Risk warning includes risk identification and assessment. In the process of risk assessment, the first step is to build a user portrait: employees dig deep into the industry to mine data and process user information and apply big data technology for business analysis and data modelling. Banks can gather product information and financial information of potential customers in terms of their portraits. Then, they can analyze multidimensional data involving different information labels, social attributes, and professional positions in order to comprehensively depict the overall image of customers. This will provide more references for the credit risk assessment. The following step is conducting big data processing: after obtaining the internal and external data according to the rules, commercial banks need to clean the data, filtering out useless repetitive information and abnormal data while keeping

the data integrity. It is also critical to analyze the variables in the data and the relationships between variables, and then transform some variables given their meanings to meet the need for risk early warning. Finally, according to the data processing model and combining both conventional methods and big data techniques, commercial banks should analyze customers' actual positions extensively, predict the probability of customer default and generate risk assessment reports [9]. For firms that are already exposed to risks, knowing their financial status can help recover funds in time. The information network formed by big data promotes the synchronization of information across various platforms, which can effectively deter borrowers who want to avoid loan repayment and drive up the costs for default companies [10].

4. Suggestions for Implementation

4.1 Improving the Ability

Applying big data technology requires commercial banks to collect and manage a huge volume of data. Data itself is the most precious resource and the key for commercial banks to take the lead in the industry in the future. Therefore, it is essential for commercial banks to consider how to outperform other competitors in data collection, establishing unique ways and channels of data acquisition. Banks can build their own database and manage information systematically in order to maintain and exploit this special asset. In addition, a commercial bank could also participate in data sharing with other banks. Cooperating with internet finance companies and big data analysis companies, commercial banks can obtain customers' credit, financial, consumption and other information in a wide range. For banks with exclusive databases, this data can serve as a supplement. For small banks that cannot build databases, cooperation can also be a fast and efficient way to take advantage of big data. Commercial banks can conveniently gather user data when providing online services, too [11].

Besides data collection, the storage of massive information is also particularly important. In managing information, banks must pay close attention to privacy, otherwise leaks of user data could lead to devastating risks and losses for commercial banks on an ethical, economic, and legal level. Ways to solve the problem include marking the information attributes, intelligently supervising the operating process, sharpening penalties for data leakage, and conducting split management of information. Among them, marking information attributes can incentivize clear terms of reference for relevant information, and help lock down the scope of the issue when information leakage occurs. What's more, different employees should be responsible for different steps in data processing. Banks ought to conduct layer management and separate the authority for information action from the data manager, thus playing the role of mutual monitoring [10].

4.2 Professional Personnel Training and Introduction

The role of technology professionals is crucial for realizing the large-scale application of big data and replacing conventional business models. Commercial banks need to develop appropriate human resources strategies and be clear about the importance of retaining competent employees. Banks should train many technical elites, including operation technicians and backstage relevant technical developers who would contribute to constantly upgrading the big data analysis model. In addition, commercial banks ought to pay attention to attracting relevant technical personnel externally. This can bring new perspectives to the team and expand the talent network, thus further improving the team's analytical ability. As part of their human resources strategy, commercial banks should define the specific positions and number of technicians needed, and assign them according to their skills. The training of talent and the way talent is recruited internally and externally must also be considered. The way to work with technical professionals is the core of the successful operation of credit businesses using big data. Especially with the continuous expansion of applications of big data analysis, commercial banks would rely more on algorithm refinement and creative thinking to gain competitive advantages. Therefore, the significance of relevant talents will only increase further.

5. Conclusion

In summary, this study examines three main applications of big data technology. Primarily, commercial banks gather large amounts of user information and depict user images to carry out targeted advertising according to user characteristics through big data technology. This information could also help develop new services and save money based on customer needs. Secondly, the traditional credit evaluation process is complicated and requires a large volume of information. Big data technology can obtain more soft information to supplement required information, simplify the process and reduce manual involvement, thus reducing cost and improving efficiency. Especially small customers would benefit from this; hence they would borrow more money from banks. Finally, using user profiles, big data can comprehensively assess the default risk of a customer and provide warnings and control well in advance. In this way, big data can offer a faster response than conventional risk control methods, and reduce losses associated with the risks. Besides, to make better use of big data technology, it is necessary to improve data acquisition, application, and management as well as the training and introduction of competent personnel. Commercial banks ought to build databases or seek cooperation with data companies and enhance staff monitoring in data management. In addition, commercial banks need to implement a sustainable human resources strategy, focusing on both internal professional training and external expert introduction. This study will help other researchers quickly understand how big data can refine the credit business and further improve the core operation capability of banks in the future. Nevertheless, this paper does not discuss the potential impacts of differences in the environment on applications of big data in the credit business. Future studies could analyze the differences based on more specific examples of bank operations to further investigate this topic. Overall, these results offer a guideline for commercial banks to create competitive advantages through the utilization of big data technology.

References

[1] Zhang Min. Applications of big data analysis in the era of big data. Automation and instrumentation, 2015, 9: 125-127.

[2] García, S., Ramírez-Gallego, S., Luengo, J. et al. Big data preprocessing: methods and prospects. Big Data Anal, 2016, 1, 9.

[3] Fan Yi. Strategy research of the credit business transformation in commercial banks from the perspective of internet finance. The financial times, 2020, 20: 91-92.

[4] Zhao Chengying. Customer relationship management of commercial banks based on big data analysis. North China Electric Power University (Beijing), 2021.

[5] Lyu Shanshan, Peng Yingying and Liu Gongchao. Research on big data analysis and application in commercial banks. China new communications, 2020, 22(01): 98.

[6] Fu Jie. The impacts of financial technology on the credit structure of commercial banks. Northwest University, 2021.

[7] Lu Huiting. Research on the innovation of commercial banks' credit business model under the background of financial technology. Academy of international trade and economic cooperation, Ministry of Commerce, 2022.

[8] Wen Xin. Research on credit risk control of commercial banks based on credit process. Hebei GEO University, 2020.

[9] Yi Sha. Research on internet credit risk warning based on big data analysis. Northwest A&F University, 2019.

[10] Zeng Qingyang. Research on applications of big data in the credit business of small and micro enterprises in commercial banks. Guangxi University, 2018.

[11] Shan Guangnian. Research on credit risk management strategies of commercial banks under the background of big data. Business and economics, 2020, 8:164-165.